

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
13 December 2001 (13.12.2001)

PCT

(10) International Publication Number
WO 01/93669 A2

(51) International Patent Classification⁷: **A01K 67/027**

SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(21) International Application Number: PCT/IB01/01199

(22) International Filing Date: 8 June 2001 (08.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(30) Priority Data:
0014009.5 8 June 2000 (08.06.2000) GB

(71) Applicant (*for all designated States except US*): **DEV-GEN NV** [BE/BE]; Technologiepark 9, B-9052 Zwijnaarde (BE).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **FEICHTINGER, Richard** [AT/BE]; Walpoortstraat 9, B-9000 Gent (BE). **BOGAERT, Thierry** [BE/BE]; Wolvendreef 26g, 8500 Kortrijk (BE).

(74) Agents: **BALDOCK, Sharon, Claire et al.**; Boulton Wade Tennant, Verulam Gardens, 70 Gray's Inn Road, London WC1X 8BT (GB).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)*
- *of inventorship (Rule 4.17(iv)) for US only*

Published:

- *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPOUND SCREENS RELATING TO INSULIN DEFICIENCY OR INSULIN RESISTANCE

(57) Abstract: The invention is concerned with use of the model organism *C. elegans* as a research tool to screen for compounds active in insulin signalling. In particular, the invention relates to improved screening methods based on release of *C. elegans* from the dauer larval state.

WO 01/93669 A2

- 1 -

**COMPOUND SCREENS RELATING TO INSULIN DEFICIENCY OR
INSULIN RESISTANCE**

The present invention is concerned with using the
5 model organism *C. elegans* as a research tool to
effectively screen compound libraries for compounds
active in insulin signalling, in particular compounds
which act downstream of the insulin receptor.
Specifically the invention relates to improved
10 screening methods based on release of *C. elegans* from
the dauer larval state.

In a particular embodiment, the invention
provides improved screening methods using *C. elegans*
carrying mutations in one or more gene(s) involved in
15 the insulin signalling pathway, such as the *Daf*-genes.
In one particular embodiment, (at least one of) said
mutation(s) is in the *daf-2* gene, which is homologous
to the insulin receptor subfamily of receptor tyrosine
kinases. On the basis of the homology between *daf-2*
20 and the insulin receptor subfamily it is proposed that
worms mutant in the *daf-2* gene may serve as models for
insulin-related diseases and disease risks, as for
example diabetes mellitus, obesity, insulin resistance
and impaired glucose tolerance (Kimura et al. 1997,
25 Science 277, 942-946).

General techniques and methodology for performing
in vivo assays using the nematode worm *Caenorhabditis*
elegans (*C.elegans*) as a model organism have been
described in the art, most notably in the following
30 applications by applicant: PCT/EP99/09710 (published
on 15 June 2000 as WO 00/34438); PCT/EP99/04718
(published on January 15, 2000 as WO/00/01846);

- 2 -

PCT/IB00/00575 (published on October 26, 2000 as WO 00/63427); PCT/IB00/00557 (published on October 26, 2000 as WO 00/63425); PCT/IB00/00558 (published on October 26, 2000 as WO 00/63426); as well as for
5 instance PCT/US98/10080 (published on 19-11-1998 as WO 98/51351), PCT/US99/13650, PCT/US99/01361 (published on 29-07-1999 as WO99/37770), and PCT/EP00/05102.

As described in these applications, one of the main advantages of assays involving the use of *C. elegans* is that such assays can be carried out in
10 multi-well plate format (with each well usually containing a sample of between 2 and 100 worms) and - also because of this - may also be carried out in an automated fashion, i.e. using suitable robotics (as
15 are described in the aforementioned applications and/or as may be commercially available). This makes assays involving the use of *C. elegans* ideally suited for screening of libraries of chemical compounds, in particular at medium to high throughput. Such
20 automated screens may for instance be used in the discovery and/or development of new compounds (e.g. small molecules) for pharmaceutical, veterinary or agrochemical/ pesticidal (e.g. insecticidal and/or nematocidal) use.

25 Some other advantages associated with the use of *C. elegans* as a model organism (e.g. in the assay techniques referred to above) include, but are not limited to:

30 - *C. elegans* has a short life-cycle of about 3 days. This not only means that these nematodes (and suitable mutants, transgenics and/or stable lines thereof) can

- 3 -

be cultivated/generated quickly and in high numbers, but also allows assays using *C.elegans* to test, in a relatively short period of time and at high throughput, the nematode worms over one or more, and
5 up to all, stages of life/development, and even over one or more generations. Also, because of this short life span, in *C.elegans* based-assays, compounds may be tested over one or more, and up to essentially all, stages of development, without any problems associated
10 with compound stability and/or (bio)availability;

- *C. elegans* is transparant, allowing -with advantage- for visual or non-visual inspection of internal organs and internal processes, and also the use of markers
15 such as fluorescent reporter proteins, even while the worms are still alive. Also, as further mentioned below, such inspection may be carried out in automated fashion using suitable equipment such as plate readers;

20 - *C.elegans* is a well-established and well-characterized model organism. For example, the genome of *C.elegans* has been fully sequenced, and also the complete lineage and cell interactions (for example of
25 synapses) are known. In addition, *C.elegans* has full diploid genetics, and is capable of both sexual reproduction (e.g. for crossing) as well as reproduction as a self-fertilizing hermaphrodite. All this may provide many advantages, not only for the use
30 of *C.elegans* in genetic and/or biological studies, but also for the use of *C.elegans* in the discovery, development and/or pharmacology of (candidate) drugs

- 4 -

for human or animal use.

Techniques for transforming, handling, cultivating, maintaining and storing (e.g. as frozen samples, which offers great practical advantages) *C. elegans* are well established in the art, for instance from the handbooks referred to below. For example, *C. elegans* may be used as one or more samples with essentially fully isogenic genotype(s).

10

Generally, in the assays described above, the nematodes are incubated in suitable vessel or container - such as a compartment or well of a multi-well plate - on a suitable medium (which may be a solid, semi-solid, viscous or liquid medium, with liquid and viscous media usually being preferred for assays in multi-well plate format). The nematodes are then contacted with the compound(s) to be tested, e.g. by adding the compound to the medium containing the worms. After a suitable incubation time (i.e. sufficient for the compound to have its effect - if any - on the nematodes), the worms are then subjected to a suitable detection technique, i.e. to measure/determine a signal that is representative for the influence of the compound(s) to be tested on the nematode worms, which may then be used as a measure for the activity of the compound(s) in the in vivo assay.

Often, in particular for automated assays, such a detection technique involves a non-visual detection method (as further described in the applications mentioned above), such as measurement of fluorescence

..

- 5 -

or another optical method, measurement of a particular marker (either associated with worms or associated with the medium) such as autonomous fluorescent proteins (AFP's) such as green fluorescent proteins (GFP's), aequorin, alkaline phosphatase, luciferase, 5 Beta-glucoronidase, Beta- lactamase, Beta-galactosidase, acetohydroxyacid, chloramphenicol acetyl transferase, horse radish peroxidase, nopaline synthase, or octapine synthase. For example, for 10 automated assays carried out in multi-well plates, so called (multi-well) "plate readers" may be used for detecting/measuring said signal.

For a further description of the above and other assay techniques involving the use of nematodes as a 15 model organism, reference is made to the prior art, such as the applications by applicant referred to above.

For general information on *C.elegans* and techniques for handling this nematode worm, reference 20 is made to the standard handbooks, such as W.B. Wood et al., "*The nematode Caenorhabditis elegans*", Cold Spring Harbor Laboratory Press (1988) and D.L. Riddle et al., "*C. ELEGANS II*", Cold Spring Harbor Laboratory Press (1997).

25 The use of *C.elegans* based assays in the field of metabolic diseases - such as obesity and diabetes - has been described in a number of applications, most notably in PCT US 98/10800 and US-A-6,225,120, which relate to the use of daf-2 mutant *C.elegans* nematodes 30 for selecting compounds active in impaired glucose tolerance and diabetes, as a model for insulin resistance.

- 6 -

One of the main objects of the present invention is to provide improved methods for the selection of compounds for the field of metabolic diseases - including but not limited to obesity, impaired glucose tolerance and type-II diabetes - which methods may be used for drug discovery, development, pharmacology and testing. In particular, it is an object of the invention to provide such improved assays as compared to the assay techniques described in PCT US 98/10800 and US-A-6,225,120.

Generally, the invention solves this problem by the use, in such assays, of nematode strains (such as m41) which have increased sensitivity of the insulin signalling pathway compared to the strains used in PCT US 98/10800 and US-A-6,225,120.

Diabetes mellitus is a major growing public health problem in both developed and developing countries. Including clinical complications it accounts for 5% of the total healthcare expenditure in Europe. Depending on the type of diabetes, current drug therapy strategy for diabetes consist of a diet supported by either application of exogenous insulin of different origin, application of drugs that increase production and/or release of endogenous insulin, enhance sensitivity of peripheral organs to insulin or mimic insulin effects. Drugs acting directly in the insulin pathway downstream of the receptor are potentially beneficial in both major types of diabetes but they are not existing today. The major drawback of currently available drugs is the body weight gain that comes on top of an existing obesity in the vast majority (80%) of patients. This

- 7 -

side effect is also the main reason why pharmacological intervention in the middle range of disease development is not as intense and aggressive, as it should be to achieve optimal efficacy. New
5 drugs that are devoid of this side effect would already reduce risk of complications by 12 to 30% (United Kingdom prospective diabetes study. Turner et al. 1998, BMJ 316: 823-828; Turner et al. 1999, JAMA 281: 2005-2012).

10 Novel glitazones, such as troglitazone, that act on nuclear receptors which regulate carbohydrate metabolism that have been launched in Japan and the US were withdrawn due to an elevated risk of liver
toxicity. Hence the medical need for well tolerated
15 orally-active anti-diabetics with mild benign side-effects remains high. A compound that directly interacts downstream the insulin receptor pathway could establish a breakthrough especially since it
could be a drug that acts both in Type I and Type II
20 diabetes. A compound that has as a clinical result an insulin sparing effect could also be of extremely high therapeutic value.

From animal studies inorganic vanadates are known to favourably combine increase in insulin sensitivity
25 and reduction of hyperlipidemia together with body weight stability or loss, but are devoid of body weight gain (Brichard and Henquin 1995, TiPS 16: 265-270). Due to unresolved toxicity issues, however, they are not available in drug formulas. Although
30 inorganic vanadium compounds are currently in clinical trial, the issue of side effects still raises doubts for this class of compounds to have to specification

- 8 -

of a drug, which has to be well tolerated in multiple doses per day for decades.

Nevertheless, the recognition of protein tyrosine phosphatase 1B as the major target of vanadates and the validation of this target as strongly increasing insulin sensitivity when inactivated in mice points towards the insulin receptor pathway as valuable for finding active compounds to ameliorate insulin resistance (Elchebly et al. 1999, Science 283: 1544-1548). PTP-1B is a negative regulator of insulin receptor tyrosine phosphorylation and kinase activity, its inactivation is raising insulin signalling with given constant insulin levels (Figure 1). The present inventors have shown that vanadates can rescue the genetic insulin resistance caused by daf-2 mutations in *Caenorhabditis elegans*, thereby validating the genetic model for insulin-deficient and insulin-resistant related disease by pharmacological means (Figure 3). Wortmannin, an inhibitor of the downstream effector phosphatidyl-inositol-3-phosphat kinase (Figure 1), further increases insulin resistance, confirming the sensitivity of the invented assay for the pathway (Figure 4). The possible known targets in the insulin-receptor pathway shown in Figure 1 are listed in table 1.

The inventors have made two key adaptations which enable them to use *C. elegans* mutant strains to effectively screen large compound libraries for activities mimicking vanadates using screens based on rescue of the phenotype dauer formation and other phenotypic traits which are caused by interventions in the insulin signalling pathway, such as, for example,

- 9 -

mutations in the insulin receptor gene homologue *daf-2*. The first adaptation is the use of *C. elegans* with a sensitized genetic background; the second adaptation is manipulation of the assay conditions such that a basal level of release from the dauer larval state is present even in the absence of test compounds. The *daf-2* gene had previously been disregarded as useful target for compound screens due to a failure of obtaining active compounds from large compound libraries (Carl Johnson, Axys pharmaceuticals, Nemapharm division, disclosed at the Cold Spring Harbor worm course). The new developments described herein overcome sensitivity problems previously encountered with screens based on *daf-2*.

In the invention, generally nematode strains are used that show sensitivity of the insulin signalling pathway.

In particular, these strains are used in assays involving the use of a dauer stage and/or dauer phenotype as a read out. These may for instance be assays based on "dauer rescue" and/or on "dauer formation/bypass" (of which dauer bypass is usually preferred, as it may avoid the problems associated with the limited uptake of the compound(s) to be tested by worms in the dauer state).

In the former type of assay, a sample of worms in the dauer state is provided, and the efficacy of the compound(s) to be tested in bringing the worms of said sample out of the dauer state is determined.

Generally, compounds with the desired activity will bring the worms out of the dauer state (i.e. to a greater degree than a reference without compound, and

- 10 -

preferably in a dose/concentration-dependant manner) and thus provide adults (i.e. more adults than without the presence of the compound(s) to be tested).

In the latter type of assay, a sample of worms
5 (in particular eggs, L1 or L2 worms, and preferably L1 worms) is kept under conditions which, without the presence of any compound(s) to be tested, would cause (most and preferably essentially all) of the worms, in the sample to enter the dauer state, and the efficacy
10 of the compound(s) to be tested in preventing the worms, under these conditions, to enter the dauer state (i.e. to bypass the dauer state) is determined. Generally, compounds with the desired activity will prevent the worms from entering the dauer state (i.e.
15 to a greater degree than a reference without compound, and preferably in a dose/concentration-dependant manner) and thus provide adults (i.e. more adults than without the presence of the compound(s) to be tested, and preferably in a dose-dependant manner). Conditions
20 such that the worm strain(s) used will enter the dauer state without the presence of the compound(s) to be tested will depend on the specific worms strain used and will be clear to the skilled person, also in view of the preferred conditions described hereinbelow.
25 Also, these conditions are preferably such that, under the conditions of the assay, a reference compound with the desired activity (such as vanadate at a concentration of between 0.5 and 2 milliMolar) will allow a measurable amount of worms to bypass the dauer
30 state (e.g. between 40 to 70%, or even more). If necessary, the results obtained with such a reference compound may also serve as a positive control or

- 11 -

comparative reference for the compound(s) to be tested.

As will be clear to the skilled person, for both the dauer rescue and the dauer bypass assays described above, and during or at the end of the assay, either the number of dauer larvae in the sample and/or the number of adults may be determined (with the sum of the number of dauer larvae and the number of adults being essentially equal to the number of worms present in the original sample). Techniques for determining the number of adults and/or dauer larvae in a sample will be clear to the skilled person and may include visual inspection of the sample (e.g. counting) as well as the automated non-visual detection techniques referred to above.

In the context of the present invention, the insulin signalling pathway may generally be described in all enzymatic conversions and other signal transduction events that are involved in (transmembrane) receptor-mediated (cellular) signal transduction in response to the (extracellular) presence insulin signals (e.g. the extracellular presence of insulin or insulin-like compounds). Some of the most important (but non-limiting) examples of the different enzymatic conversions involved in said signalling have already been mentioned hereinabove.

By "*sensitivity of the insulin signalling pathway*" is generally meant that

1) the nematode shows one or more biological response(s) to the presence of an insulin, to the presence of an insulin-like compound, and/or to the presence of a compound that can provide and/or or

- 12 -

mimic a biological response similar to the biological response(s) provided by insulin or the insulin-like molecules (which three categories are also collectively referred to herein as "*insulin-like signals*"); and that

5 2) said one or more biological responses change when (the amount of) the compound(s) to which the nematode is exposed (and/or with which said nematode comes into contact) changes or is altered
10 (for instance, due to a change in the concentration of said insulin like signal in the medium.

The biological response may be any response or combination of responses, such as one or more changes in physiology, biochemistry, development, behaviour, exitation, or other phenotypical properties.

15 In one particularly preferred embodiment, these may essentially be one or more of the biological responses that are (also) obtained upon (over)expression of insulin the nematode.

20 One particularly suited biological response may be the dauer-behaviour, e.g. the entry, exit, rescue or bypass of the dauer state, and/or other phenotypical properties that result from and/or are associated with the so-called dauer decision.

25 In the invention, (one or more strains of) nematodes are used that show increased sensitivity of the insulin pathway, compared to at least the wildtype, and preferably also compared to the reference strain CB1370 (containing the *daf-2* reference mutation *el370*. This strain is publicly
30 available, for example from the Caenorhabditis Genetics Center (CGC), Minnesota, USA).

- 13 -

By "increased sensitivity of the insulin signalling pathway" is generally meant that the change in the biological response of the nematode (as described above) to a change in (the concentration of) the insulin-type signal is greater than the change that is obtained with the wildtype and/or CB1370 (i.e. for the same change in (the concentration of) the insulin-type signal).

For example, when a change in (e.g. an increase or reduction of) the concentration of an insulin-type signal gives, for the wildtype and/or CB1370, a change in (e.g. an increase or reduction of) the biological response of by a factor of x , than the same change will give, for a strain suitable for use in the invention, a change in the same biological response of more than x (e.g. 1.05 times x , preferably 1.1 times x , more preferably 1.5 times x or even 2 times x or 10 times x , depending on the biological response, the insulin-type signal, the change in concentration, and the specific strain(s) used). In case there is no change observed in wildtype and/or the reference strain CB1370, any change observed determines a strain to be of "increased sensitivity to a insulin-type signal".

For example, an "insulin-type signal" as used herein may be:

- an insulin or insulin-like molecule (e.g. from any suitable source, including but not limited to nematodes, humans or other animals), for which reference is made to PCT/US99/08522, published as WO99/54436 on 28.10.99; Genes & Development 15:672-686,2001;

- 14 -

- a vanadate or a vanadate-type compound, such as sodium orthovanadate;
- a PTB-1B inhibitor such as described in Journal of Medicinal Chemistry 43:1293-1310, 25.02.2000, for
5 example compound 66;
- wortmannin or a wortmannin-type compound, such as LY 294002 or other PI3-kinase inhibitors.

In this respect, it should be noted that an increase in the concentration of an insulin-type
10 signal may provide an increase in the biological response (in which said increase will be more pronounced for the strain of the invention than for the wildtype and/or for CB1370), or may provide a decrease in the biological response (in which said
15 decrease will be more pronounced for the strain of the invention than for the wildtype and/or for CB1370). For example, an increase in the concentration of a wortmannin will provide an increase in the biological response (for example more dauer), which will be even
20 more pronounced for the strains of the invention (e.g. even more dauer compared to wildtype/CB1370 per increased concentration of wortmannin), whereas an increase in the concentration of a vanadate will provide a decrease in the biological response (for
25 example less dauer), which will be even more pronounced for the strains of the invention (e.g. even less dauer compared to wildtype/CB1370 per increased concentration of vanadate). In case the number of nematodes grown up, i.e. non-dauer, are counted,
30 positive (i.e. increased) and negative (i.e. decreased) biological response are reversed into each other. Both types of insulin-type signals may be used

- 15 -

for to determine whether a specific nematode strain has "*increased sensitivity of the insulin signalling pathway*" compared to wildtype and/or CB1370, and which may be used within the scope of the present invention.

5 Preferably, the insulin-type signal that is used to determine whether a specific nematode strain has "*increased sensitivity of the insulin signalling pathway*" is a vanadate-type compound. The vanadate may be used as a free base or as a suitable water-soluble
10 salt, such as sodium orthovanadate. Preferably, the vanadate is used in an amount of between 0.01 and 100 millimolar, more preferably between 0.1 and 10 millimolar, such as 0.5 millimolar or 2.0 millimolar.

 Some specific conditions under which vanadates
15 may be used to determine whether a specific nematode strain has "*increased sensitivity of the insulin signalling pathway*" will be further described below.

 Thus, as will be clear from the above, the
20 "insulin-type factor(s)" described above may be used to determine whether a strain has increased sensitivity of the insulin signalling pathway (i.e. compared to the wildtype and/or CB1370) and thus may be used within the scope of the invention.

 Generally, such a nematode strain useful in the
25 invention will have "*increased sensitivity of the insulin signalling pathway*" due to a mutation and/or an other genetically determined factor that provides such increased sensitivity. Such strains will also be referred to below as having a "sensitized genetic
30 background", and some preferred examples thereof, such as DR1564 and CB1368, will be further described below.

- 16 -

However, it is also within the scope of the invention to provide the strain(s) used with "increased sensitivity of the insulin signalling pathway" by other means, such as exposure to
5 pheromones which increase such sensitivity, by gene suppression techniques such as RNAi, and/or by growing/cultivating the nematodes in the presence of an inducing or suppressing factor (such as population density, food concentration and temperature).

10 In particular, the nematode strain used may be a weak *Daf* mutant (i.e. a mutation abnormal in dauer formation), in particular a *Daf* mutant that is weaker than the reference strain CB1370. For instance, it may be a *age-1* mutant, or one of the other *daf* mutants
15 mentioned herein.

In particular, the nematode strain used may be a weak *daf-2* mutant, in particular a *daf-2* mutant that is weaker than the reference strain CB1370.

For instance, the reference strain used may be
20 have a Class-I mutation (as mentioned in Gems et al., supra), a mutation which provides a phenotype similar to - and preferably essentially the same as - a Class-I mutation, and/or a(nother) mutation in the ligand binding domain, such that the mutated receptor still
25 has an active kinase domain, but the sensitivity to insulin-like signalling is impaired. However, in its broadest scope, the invention is not limited thereto, and other mutations may also be present, including Class II mutations, as long as the strain having the
30 mutation still has increased sensitivity of the insulin signalling pathway, compared to the wildtype and/or the reference strain *C. elegans* CB1370.

- 17 -

It is also possible, in the assays of the invention, to use two or more different strains, e.g. one or more which have increased sensitivity of the insulin signalling pathway, and/or one or more references, e.g. wildtype or CB1370.

In one preferred, but non-limiting aspect of the invention, the sensitivity of the insulin signalling pathway of the nematode strain used may be expressed in terms of the "*Insulin Sensitivity Value*" (ISV), which may be determined in the following manner:

A sample of nematode worms (preferably in the L1 stage) is incubated for between 48 and 96 hours (preferably about 72 hours) separately with and without an insulin-type signal (preferably a vanadate-type compound), at a temperature of between 20 and 25°C (such as 20, 21, 22, 23, 24 or 25°C), in the presence of a suitable source of food (such as bacteria, e.g. between 0.05 and 0.5 % w/v, preferably about 0,125 % w/v), and using a suitable medium (such as S-buffer, M9 or one of the media described in the applications referred to above, and preferably S-buffer).

After incubation, for both the sample with the insulin-type signal and the sample without the insulin-type signal compound, the number of worms in the sample that enter into the dauer state is determined, as a percentage of the number of worms in the original sample, i.e. as follows:

- 1) for the sample without the insulin-type signal:
([the number of worms that enter the dauer state without insulin-type signal] divided by [the

- 18 -

total number of L1 worms in the original sample))
times [100%].

This percentage is herein referred to as "Percentage A".

5

2) for the sample with the insulin-type signal:

([the number of worms that enter the dauer state
with the insulin-type signal] divided by [the,
total number of L1 worms in the original sample])
times [100%].

10

This percentage is herein referred to as "Percentage B".

The Insulin Sensitivity Value may then be
expressed as the absolute difference between

15

"Percentage A" and "Percentage B" (i.e. as absolute
value of ["Percentage A" minus "Percentage B"]).

As the ISV is calculated as a difference between
two percentages A and B, the ISV itself will be a
percentage (for instance, when Percentage A is 90%,
and percentage B is 10%, the ISV will be $90\% - 10\% =$
80%), and always positive as the absolute value is
calculated (for instance, when Percentage A is 10% and
Percentage B is 90%, the ISV will be $|10\% - 90\%| = |-$
80%| = 80%).

20

25

In the invention, the nematode strain used
preferably has an ISV that is greater than the ISV for
CB1370. In particular, the nematode strain used may be
such that its ISV is more than 1% greater, preferably
more than 5% greater, more preferably more than 10%
greater, even more preferably more than 20% greater
than the ISV for CB1370 (e.g. calculated as the
absolute difference between the ISV for the strain

30

- 19 -

used and the ISV for CB1370, e.g. [ISV strain used] minus [ISV CB1370]).

For example, depending upon the specific conditions of the test, CB1370 will usually have an ISV of <20%, more usually <10%, and often <5% (in essence, this means that under the conditions of the test, for CB1370, there is little no difference between the presence and the absence of the insulin type signal). The ISV for wildtype will usually be even lower than the ISV for CB1370.

For the strain used in the invention, under the same conditions of the test, the ISV will usually be >30 %, and is preferably >40%, and is even more preferably >50%. (in essence, this means that under the conditions of the test, for the strain used, the difference between the presence and the absence of the insulin-type signal is preferably (much) larger than for CB1370).

Preferably, the ISV is determined using a vanadate-type compound such as sodium orthovanadate, although the invention in its broadest sense is not limited thereto.

Thus, by determining the ISV in the manner outlined above, it can be determined whether a strain has increased sensitivity of the insulin signalling pathway, compared to the wild-type and/or the reference strain CB1370.

Generally, the invention is based on the insight that such nematode strains having increased sensitivity of the insulin signalling pathway can be used with advantage to provide improved methods for the selection of compounds for the field of metabolic

- 20 -

diseases, in particular compared to the assay techniques described in PCT US 98/10800 and US-A-6,225,120. As mentioned above, these methods may be used for drug discovery, development and pharmacology, for instance to discover and/or develop new small molecules and/or small peptides suitable for use in preventing or treating metabolic diseases in human or vertebrates (such as mammals).

For the purposes of the present disclosure, a "small molecule" generally means a molecular entity with a molecular weight of less than 1500, preferably less than 1000. This may for example be an organic, inorganic or organometallic molecule, which may also be in the form of a suitable salt, such as a water-soluble salt.

The term "small molecule" also covers complexes, chelates and similar molecular entities, as long as their (total) molecular weight is in the range indicated above.

In a preferred embodiment, such a "small molecule" has been designed according, and/or meets the criteria of, at least one, preferably at least any two, more preferably at least any three, and up to all of the so-called Lipinski rules for drug likeness prediction (vide Lipinski et al., Advanced Drug Delivery Reviews 23 (1997), pages 3-25). As is known in the art, small molecules which meet these criteria are particularly suited (as starting points) for the (design and/or) development of drugs (e.g) for human use, e.g. for use in (the design and/or compiling of) chemical libraries for (high throughput screening), (as starting points for) hits-to-leads chemistry,

- 21 -

and/or (as starting points for) lead development.

In a preferred embodiment, such a "small molecule" has been designed according, and/or meets the criteria of, at least one, preferably at least any
5 two, more preferably at least any three, and up to all of the so-called Lipinski rules for rational drug design (vide Lipinski et al., Advanced Drug Delivery Reviews 23 (1997), pages 3-25). As is known in the art, small molecules which meet these criteria are
10 particularly suited (as starting points for) the design and/or development of drugs (e.g) for human use

Also, for these purposes, the design of such small molecules (as well as the design of libraries consisting of such small molecules) preferably also
15 takes into account the presence of pharmacophore points, for example according to the methods described by I. Muegge et al., J. Med. Chem. 44, 12 (2001), pages 1-6 and the documents cited herein.

The term "small peptide " generally covers
20 (oligo)peptides that contain a total of between 2 and 35, such as for example between 3 and 25, amino acids (e.g. in one or more connected chains, and preferably a single chain). It will be clear that some of these small peptides will also be included in the term small
25 molecule as used herein, depending on their molecular weight.

Thus, the methods of the invention may in particular be used to test and/or screen (libraries of) such small molecules and/or peptides, in the
30 manner as further outlined herein.

Thus, in one aspect, the invention relates to the use of at least one nematode worm which has an

- 22 -

increased sensitivity of the insulin signalling pathway (compared to the wildtype and/or the reference strain CB1370), in an assay for the identification of a compound, such as a small molecule and/or a small peptide, which is capable of modulating insulin signalling pathways (for example in *C. elegans* and/or vertebrates, such as humans and/or other mammals), more generally of altering and/or effecting the biological response to insulin signalling, and even more generally for use in (the preparation of compositions for) the prevention and/or treatment of metabolic diseases or disorders (as mentioned above), in vertebrates such as humans or other mammals.

In addition to the identification of small molecules and/or small peptides, according to the inventions, the nematode worms with an increased sensitivity of the insulin signalling pathway may also be used for determining the influence or effect of gene suppression (e.g. by RNAi techniques), and of specific or non-specific mutations (e.g. due to non-specific or (site-)specific mutagenesis).

Preferably, the nematode worm with increased sensitivity of the insulin signalling pathway has a sensitized genetic background (compared to the wildtype and/or the reference strain CB1370), as defined above.

Even more preferably, the nematode worm with increased sensitivity of the insulin signalling pathway (e.g. a sensitized genetic background) has an ISV which is greater than the ISV for wildtype and/or CB1370, and even more preferably an ISV as defined above.

- 23 -

Some preferred, but non limited examples of suitable *C. elegans* strains include, but are not limited to: DR1564: *daf-2(m41)*, CB1368: *daf-2(e1368)* and some of the (other) strains mentioned in Gems et al., supra. Other suitable strains will be clear to the skilled person, based upon the disclosure herein.

The most preferred nematode strain is DR1564: *daf-2(m41)*.

The sample of nematodes may comprise any suitable number of worms, depending on the size of the container/vessel used. Usually, the sample will comprise between 2 and 500, in preferably between 3 and 300, more preferably between 5 and 200, even more preferably between 10 and 100 nematodes. When the assay is carried out in multi-well plate format, each well usually contains between 15 and 75 worms, such as 20 to 50 worms. Although not preferred, it is not excluded that a sample may consist of a single worm.

Usually, each such individual sample of worms will consist of worms that - at least at the start of the assay - are essentially the same, in that they are of the same strain, in that they contain the same mutation(s), in that they are essentially of an isogenic genotype, in that they show essentially the same phenotype(s), in that they are essentially "synchronised" (i.e. at essentially the same stage of development, such as L1 or dauer. It should however be noted that this stage of development may - and usually will - change during the course of the assay, and not for all worms in the sample at the same rate and/or in the same way), in that they have been grown/cultivated in essentially the same way, and/or in that they have

- 24 -

been grown under and/or exposed to essentially the same conditions, factors or compounds, including but not limited to pheromones, gene suppression (such as by RNAi), gene- or pathway-inducing factors or (small) molecules, and/or gene- or pathway-inhibiting factors or (small) molecules. However, in its broadest sense, the invention is not limited thereto.

The medium may further contain all factors, compounds and/or nutrients required to carry out the assay and/or required for the survival, maintenance and/or growth of the worms. For this, reference is again made to the prior art, such as the applications and handbooks referred to above. In one specific embodiment, the medium may also contain a suitable source of food for the worms - such as bacteria (for example a suitable strain of *E. coli*) - in a suitable amount.

In the method of the invention, the sample of nematodes can be kept - e.g. maintained, grown or incubated - in any suitable vessel or container, but is preferably kept in a well of a multi-well plate, such as standard 6, 24, 48, 96, 384, 1536, or 3072 well-plates (in which each well of the multi-well plate may contain a separate sample of worms, which may be the same or different). Such plates and general techniques and apparatus for maintaining/ handling nematode worms in such multi-well plate format are well known in the art, for instance from the applications mentioned hereinabove.

The sample of nematodes may be kept in or on any suitable medium - including but not limited to solid and semi-solid media - but is preferably kept in a

- 25 -

suitable liquid or viscous medium (e.g. with a viscosity at the temperature of the assay that is equal to a greater than the viscosity of M9 medium, as measured by a suitable technique, such as an
5 Ubbelohde, Ostwald and/or Brookfield viscosimeter).

Generally, suitable media for growing/maintaining nematode worms will be clear to the skilled person, and include for example the media generally used in the art, such as M9, S-buffer, and/or the further
10 media described in the applications and handbooks mentioned hereinabove.

Preferably, the assays of the invention are based on the dauer phenotype as a biological read out, e.g. the entry into, the bypass of and/or the rescue from
15 the dauer state, and/or any other property which results from and/or is associated with the so-called dauer decision.

For instance, an assay based upon entry into/bypass of the dauer state may comprise the
20 following steps:

- a) providing a sample of nematode worms (preferably eggs, L1 or L2 worms, and most preferably L1 worms);
- b) keeping said sample under conditions such, without
25 the presence of any compound(s) to be tested, at least 50%, and preferably at least 60 %, and more preferably at least 70 %, even more preferably at least 80 %, such as 85-100% of the nematodes present in said sample would enter the dauer state
30 (at least during the time used for the assay, such as at least 1 day, for example 2-4 days - e.g. about 72 hours - as further described below);

- 26 -

- c) exposing the sample to the compound(s) to be tested;
- d) measuring either the number of worms that enter the dauer state, and/or measuring the number of worms that grow into adults.

Preferably, in such an assay, the conditions used in step b) are such that, in the presence of a reference compound (such as a vanadate compound, e.g. sodium orthovanadate) at a suitable concentration (such as between 0.5 and 2 milliMolar, which is particularly suited for vanadate), the amount of worms that enter the dauer state is at least 10% less (i.e. lower in absolute difference of percentages as also referred to above), preferably at least 20% less, more preferably at least 30% less, than the amount of worms that enter the dauer state without the presence of any such reference compound (at least during the time used for the assay, such as at least 1 day, for example 2-4 days - e.g. about 72 hours - as further described below).

For instance, the conditions used in step b) may be such that, in the presence of a reference compound (such as a vanadate compound, e.g. sodium orthovanadate) at a suitable concentration (such as between 0.5 and 2 milliMolar, which is particularly suited for vanadate), the amount of worms that enter the dauer state is less than 50%, preferably less than 40%, even more preferably less than 30% (at least during the time used for the assay, such as at least 1 day, for example 2-4 days - e.g. about 72 hours - as further described below, and depending on the amount of worms that would enter the dauer state without the

- 27 -

presence of the reference), although the invention in its broadest sense is not limited thereto.

An assay based upon rescue from the dauer state
5 may comprise the following steps:

- a) providing a sample of nematode worms in the dauer state;
- b) keeping said sample under conditions such that, without the presence of any compound to be
10 tested, least 50%, and preferably at least 60 %, and more preferably at least 70 %, even more preferably at least 80 %, such as 85-100% of the nematodes present in said sample would remain in the dauer state (at least for the time
15 of the assay, such as between 1 and 96 hrs, such as between 12 and 72 hours, such as about 24-48 hours);
- c) exposing the sample to the compound(s) to be tested;
- 20 d) measuring either the number of worms that remain in the dauer state, and/or measuring the number of worms that go out of the dauer state (e.g. become adults).

Preferably, in such an assay, the conditions used
25 in step b) are such that, in the presence of a reference compound (such as a vanadate compound, e.g. sodium orthovanadate) at a suitable concentration (such as between 0.5 and 2 milliMolar, which is particularly suited for vanadate), the amount of worms
30 that remain in the dauer state is at least 10% less (i.e. lower in absolute difference of percentages as also referred to above), preferably at least 20% less,

- 28 -

more preferably at least 30% less, than the amount of worms that remain in the dauer state without the presence of any such reference compound (at least during the time used for the assay, such as between 1
5 and 96 hrs, such as between 12 and 72 hours, such as about 24-48 hours).

For instance, the conditions used in step b) may be such that, (such as a vanadate compound, e.g.
10 sodium orthovanadate) at a suitable concentration (such as between 0.5 and 2 millimolar, which is particularly suited for vanadate), the amount of worms that remain in the dauer state is less than 50%, preferably less than 40%, even more preferably less
15 than 30% (at least during the time used for the assay, such as between 1 and 96 hrs, such as between 12 and 72 hours, such as about 24-48 hours, and depending on the amount of worms that would remain in the dauer state without the presence of the reference), although
20 the invention in its broadest sense is not limited thereto.

Techniques for distinguishing, in a sample, and preferably in an automated and/or multi-well plate
25 format, the number of adults and/or the number of dauers will be clear to the skilled person and may include visual/manual techniques, and/or the non-visual detection techniques described in the applications referred to above.

30 In the assays of the invention, each individual sample of nematode worms will generally be exposed to a single compound to be tested, at a single

- 29 -

concentration; with different samples (e.g. as present in the different wells of the multi-well plate used) being exposed either to different concentrations of the same compound (e.g. to establish a dose response curve for said compound), to one or more different compounds (which may for instance be part of a chemical library and/or of a chemical class or series, such as a series of closely related structural analogues), or both (e.g. to the same and/or different compounds at different concentrations).

It is also within the scope of the invention to expose the (sample of) nematodes to two or more compounds - at essentially the same time or sequentially (e.g. with an intermediate washing step) - for example to determine whether the two compounds have an effect which is the same or different from both the compounds separately (e.g. to provide a synergistic effect or an inhibitory or competitive effect).

Furthermore, it is within the scope of the invention to use one or more reference samples, e.g. samples without any compound(s) present, and/or with a predetermined amount of a reference compound. The invention also includes the use, in an assay, of two or more samples of nematode worms of different strains, e.g. to compare (the effect of the compound(s) to be tested on) the different strains, in which said different strains may also be reference strains, such as wildtype, N2 or Hawaiian.

In a preferred embodiment, an assay based on dauer entry/bypass is carried out in a multiwell plate format, under the following conditions:

- 30 -

- use of a sample of between 2 and 100, preferably between 10 and 80, more preferably between 15 and 60 worms, such as 20 or 50 worms, preferably eggs, L1 or L2, most preferably L1.
- 5 - a temperature of between 10°C and 30 °C, preferably between 20°C and 27 °C, such as 21, 22, 23, 24, 25 or 26°C, depending on the specific strain used.
For example, for DR1564: *daf-2(m41)*, usually a
10 temperature of about 21, 22, 23, 24 °C will be preferred, with a temperature of between 21 and 22°C being particularly preferred.
For CB1368: *daf-2(e1368)*, usually a temperature of
24, 25 or 26°C will be preferred, with 25°C being
15 particularly preferred.
- a concentration of the compound(s) to be tested of between 0.1 nanomolar and 100 milimolar, preferably between 1 nanomolar and 10 milimolar, more preferably between 1 micromolar and 200
20 micromolar, such as about 20 micromolar. The compound may be taken up by the nematodes in any suitable manner, such as by drinking, soaking, via the gastrointestinal tract (e.g. the gut), via the cuticle (e.g. by diffusion or an active transport
25 mechanism), and/or via openings in the cuticle, such as amphid sensory neurons. Generally, the compound will be mixed with or otherwise incorporated into the medium used;
- a time of contact with the compound(s) to be
30 tested of between 0.1 minute and 100 hours, preferably between 1 minute and 90 hours, such as

- 31 -

- about 1 hour to 72 hours. For instance, the sample of nematodes may be contacted with the compound(s) to be tested for only a brief period of time, e.g. between 1 minute and 2 hours, such as between 20 minutes and 1.5 hours, upon which the sample of nematodes may be washed and further cultivated on fresh medium (i.e. without compound), or the sample of nematodes may be contacted with the compound(s) to be tested for essentially the entire duration of the assay (e.g. for 1-3 days or more). For assays involving (the bypass of) dauer formation (e.g. starting from L1), the time of contact will generally encompass two or more stages of development, and most preferably be between 1 and 4 days, such as about 2-3 days (e.g. 48 to 72 hours).
- a (total) time of incubation of the sample of between 0.1 minute and 100 hours, preferably between 1 minute and 90 hours, such as about 1 hour to 72 hours. For assays involving dauer entry/bypass (e.g. starting from L1), the total incubation time will generally encompass two or more stages of development, and most preferably be between 1 and 4 days, such as about 2-3 days (e.g. 48 to 72 hours);
 - the use of a liquid or viscous medium (in which viscous is as defined above), such as S-buffer, M9 or one of the other media referred to in the patent applications mentioned above (as referred to above), with S-buffer being particularly preferred.
 - The presence of a suitable source of food - for example bacteria such as *E. coli* - in a suitable

- 32 -

amount, e.g. between 0.001 and 10 % (w/v), preferably between 0.01 and 1%, more preferably between 0.1 and 0.2 %, such as about 0.125 % w/v, based on the total medium.

5 Conditions for assays based on dauer rescue are further described below and/or in PCT US 98/10800 and US-A-6,225,120.

Although the conditions described above are particularly preferred, more generally, according to
10 the invention, the nematode strains with increased sensitivity of the insulin signalling pathway (as further defined above) may be used with advantage in any *C. elegans*-based assay technique involving and/or relating to insulin-signalling, insulin signal
15 transduction, biological responses to insulin and/or insulin-type compounds, and/or the insulin pathway. These assays may be based on any suitable phenotypical read out, including but not limited to dauer entry, bypass and/or rescue as described above.

20 Therefore, in accordance with one aspect of the invention, there is provided a method for the identification of a compound which is capable of modulating insulin signalling pathways, which method comprises:

25 providing *C. elegans* larvae of a strain of sensitized genetic background to the insulin signalling pathway;

 contacting said larvae with a test compound in growth favouring conditions, i.e. including food; and

30 screening for growth to adulthood, i.e. bypass of or release from the dauer larval state.

 A "sensitized genetic background" may be defined

- 33 -

herein by comparison to the reference *daf-2* allele, *e1370* (Figure 2 is a print of the acedb database entry on *daf-2*). The term "sensitized genetic background" encompasses *C. elegans* strains which exhibits greater
5 sensitivity to test compounds than the *daf-2(e1370)* allele.

The method of the invention is suitable for use with essentially any *C. elegans* strain which exhibits a dauer phenotype as a result of defect, for example a
10 mutation, in a gene encoding a component of the insulin signalling pathway or other intervention affecting the insulin signalling pathway and which exhibits a "sensitized genetic background" as compared to the *daf-2(e1370)* mutant.

15 In a preferred embodiment the method of the invention may be carried out using *C. elegans* strain DR1564 containing the *daf-2(m41)* mutation which exhibit a dauer-constitutive phenotype. Use of strains carrying this allele in compound screens based
20 on bypass of/rescue from dauer is illustrated in the accompanying Examples. Table 6 compares the activity of 94 compounds, which were found to be positive in a primary screen of 8,000 compounds using DR1564: *daf-2(m41)*, as part of Example 1, in a retest on the
25 *m41* allele bearing strain DR1564 and on the *daf-2* alleles bearing strains CB1368: *daf-2(e1368)* and *daf-2(e1370)*. DR1564: *daf-2(m41)* was found to be more sensitive to compound activities than CB1368: *daf-2(e1368)*, with 56% and 27% confirmation rate,
30 respectively. The strain CB1370 containing the *daf-2* reference allele *e1370* could not be rescued by any of the 94 compounds.

- 34 -

Other sensitized backgrounds in addition to *daf-2(m41)* may be used in accordance with the invention. Since both *m41* and *e1368* belong to class I alleles in the classification of Gems et al. 1998, Genetics 150: 129-155, while *e1370* belongs to class II, it is likely that other class I alleles are also useful as sensitized genetic background. Typically class I alleles are mutations in the ligand binding domain, and class II mutations are located in the kinase domain. The precise molecular lesion of *m41* is unknown.

Other *C. elegans* strains with sensitized genetic backgrounds which may be used in accordance with the invention include strains exhibiting a dauer phenotype which comprise loss of function or reduction of function mutations in genes downstream of the insulin receptor (*daf-2*). A particular example is the *age-1* mutation, a mutation in the catalytic subunit of the PI3-kinase (see Figure 1 and table 1). While gain of function alleles of *akt-1* or *pdk-1* are not able to rescue *daf-2(e1370)*, they do rescue *age-1* mutations (Paradis and Ruvkun 1998, Genes & Dev 12:2488-2489, Paradis and Ruvkun 1999, Genes & Dev 13:1438-1452).

While there are no mutations known in the regulatory subunit of the PI3-kinase (located on the yac clones Y119C1 and Y110A7), knock-out mutations in these genes may be generated by methods known by the art (Zwaal et al. 1993, PNAS 90: 7431-35; Liu et al. 1999, Genome Research 9:859-867). Other suitable strains carry loss of function mutations in the genes encoding AKT protein kinases. Since there are two redundantly acting AKT protein kinases (Paradis and

- 35 -

Ruvkun 1998, Genes & Dev 12:2488-2489), a double mutation of knock-outs of both *akt-1* and *akt-2* may be to be constructed by simple crossing. Another potential useful mutation is the loss of function
5 mutation in *pdk-1(sa680)*, as described in Paradis and Ruvkun 1999, above cit.

In a further embodiment of the method of the invention, a *C. elegans* strain having a sensitized genetic background may be obtained by inhibiting
10 proteins of the insulin-receptor pathway using specific inhibitor compounds. In particular, inhibitors of the PI3-kinase are known, such as Wortmannin and LY294002. Barbar et al. 1999, Neurobiol Aging 20:513-519 demonstrate the activity of LY294002
15 in inducing dauer formation. The inventors own experiments also illustrate the activity of Wortmannin (Figure 4).

RNAi inhibition is still another method of generating *C. elegans* strains with loss of function
20 phenotypes suitable for use in the method of the invention. Methods of inhibiting expression of specific genes in *C. elegans* using RNAi are well known in the art and described, for example by Fire et al., Nature 391:801-811 (1998); Timmins and Fire, Nature
25 395:854 (1998) and Plaetinck et al., WO 00/01846. Most preferred are the techniques described in WO 00/01846 which use special bacterial strains as food source to obtain double stranded RNA inhibition.

In yet another embodiment of the present
30 invention, sensitized strains may be used which comprise gain of function mutations of *daf-18* or *daf-16* or of the *C. elegans* homologs of PTP-1B or

- 36 -

SHIP2. Generation of gain of function mutations of serine or threonine phosphorylation sites, as disclosed for *daf-16* by Paradis and Ruvkun 1998, above cit., and by Kops et al. 1999, Nature 398: 630-634, is straightforward for researchers experienced in the state of the art, as demonstrated by Nakae et al. 2000, EMBO 19: 989-996 for FKHR, a human homologue of *daf-16*.

Yet another sensitized genetic background may be derived by using mutants defective in perception of environmental signals that regulate insulin signalling, such as pheromone, food and temperature signals, or mutations in the neural processing of said signals, or mutations in the secretion of insulin-like molecules or in one of the genes encoding for an insulin-like molecule. In a preferred embodiment *tph-1(mg280)* is used, a mutant deficient in tryptophan hydroxylase, necessary for serotonin biosynthesis. *C. elegans* worms with this mutation accumulate large stores of fat and to some extent form dauer larvae because of inability to process the food sensation, together with impaired temperature sensation (Sze et al. 2000, Nature 403: 560-564). Other suitable sensitized genetic backgrounds comprise *daf-c* mutations in *daf-1*, *daf-4*, *daf-7*, *daf-8*, *daf-11*, *daf-14*, *daf-21*, *daf-19* or *daf-28*. Furthermore, dominant activation mutations in neuronal G proteins, as described by Zwaal et al. 1997, Genetics 145: 715-727, may also serve as sensitized background.

Several synthetic dauer forming mutations are known, which enhance other genetic backgrounds to form dauer mutations. One specific example, the double

- 37 -

unc-64(e246); *unc-31*(e928), is given by Ailion et al. 1999, PNAS 96, 7394-7397. Since *unc-64* encodes for a homolog of syntaxin, a protein involved in synaptic transmission and other types of Ca^{2+} -regulated secretion and *unc-31* encodes for a homolog of CAPS, Ca^{2+} -dependent activator protein for secretion and insulin release in pancreatic β cells is determined by Ca^{2+} -regulated secretion the simplest model is that the *Daf-c* phenotype of the double mutation is caused by a shut down of release of either insulin like molecules themselves or of neurotransmitters that stimulate insulin release (Ailion et al. 1999, PNAS 96, 7394-7397).

Sensitized worm strains which comprise any combination of two or more synthetic dauer formation mutations amongst each other, or in combination with dauer constitutive mutations, as examples are provided above, or any combination of dauer constitutive mutations with each other may be used in the method of the invention. An example can be drawn from Ogg et al. 1997, Nature 389: 994-999, where a *daf-2; daf-1* double mutant induces dauer formation at temperatures far below temperatures necessary for each of the single mutation to induce dauer formation.

The disclosed screening method is based on bypass of/release from the dauer larval state. There are several different ways in which to screen for bypass of/release from the dauer state which may be used in accordance with the invention, as described below. Furthermore, it is possible to use phenotypes of *Daf* genes other than dauer, including but limited to, fat storage, regulation of metabolic enzymes or

- 38 -

stress resistance pathways or any other biochemically, transcriptionally or posttranscriptionally regulated effect that is measurable as the basis of an assay read-out in accordance with the invention.

5

In accordance with a second aspect the invention also provides a method for the identification of a compound which is capable of modulating insulin signalling pathways, which method comprises:

10 providing *C. elegans* larvae of a strain of sensitized genetic background to the insulin signalling pathway;

contacting said larvae with a test compound in growth favouring conditions, i.e. including food; and
15 screening for growth to adulthood, i.e. bypass of or release from the dauer larval state, wherein conditions of assay are selected such that a basal level of bypass of or release from the dauer larval state is observed in the absence of the test compound.

20 The second aspect of the present invention comprises of a sensitized assay condition, in contrary to tight screening conditions usually performed in screens to isolate genetic suppressors of *daf-2*, e.g. *daf-16* alleles (Riddle et al. 1981, Nature
25 290:668-671; Gottlieb & Ruvkun 1994, Genetics 137: 107-120).

The inventors provide a method of setting the assay conditions in way that a basal level of release from the dauer larval state is already present in
30 controls. The basal level of release from the dauer larval state may for example be measured by counting the number of worms growing beyond the dauer stage in

- 39 -

a sufficiently large number of control wells
(containing the solvent alone but no test compounds).

The basal level of release from the dauer larval
state will preferably be between 0.1% and 60% rescue,
5 more preferably between 1% and 50% rescue and most
preferably between 2% and 40% rescue, such as 10% to
20% rescue. While the minimal number of growing worms
or residual activity is derived from sensitizing the
assay conditions, the maximal number is derived from
10 experience to optimise signal to noise ratio.

Although in a preferred embodiment the method of
the invention uses the temperature sensitivity of *daf-*
2 mutations, such as *m41*, to sensitize assay
conditions, any set of conditions that sensitize the
15 assay over the strict genetic screen conditions is
within the scope of the invention, in particular
conditions that show growth between 0.1% and 60%,
preferentially between 1% and 50%, most preferentially
between 2% and 40%, such as 10% to 20%, in cases where
20 the readout of the assay is related to bypass of or
release from the dauer-constitutive phenotype.

Another embodiment of the invention uses genetic
means to sensitize assay conditions to the desired
basal level of release from the dauer larval state.
25 For example Ogg & Ruvkun (1998), Mol. Cell 2: 887-893,
disclose a double mutation *daf-2; daf-18*, which gives
rescue (L4 and adults) at a level of 2.2%. In
addition, mutations known as *Daf-d* for dauer
defective, especially weak mutations, can be used in
30 the present invention. Also gain of function
mutations, as there are known *pdk-1(mg142)*, (Paradis
and Ruvkun 1999, Genes & Dev 13:1438-1452) and

- 40 -

akt-1(mgl144), (Paradis and Ruvkun 1998, Genes & Dev 12:2488-2489), can be used to rescue from dauer formation to a certain percentage. Furthermore, gain of function, in particular at phosphorylation sites, or loss of function mutations can be generated by methods known in the art (see citations in the section further above).

Also suitable for use in the method of the invention are *C. elegans* strains which comprise a mutation in a gene downstream of the insulin receptor in the insulin signalling pathway which leads to a reduction in the function of the product of the mutated gene but not a complete loss of function. Residual activity of the product encoded by the gene mutated in such strains may be sufficient to confer a basal level of release from the dauer larval state.

Another embodiment of the invention comprises the incomplete loss of function typically seen with RNAi experiments. Since the disclosed methods rely on growth of worms in presence of *E. coli*, methods of obtaining RNA inhibition via feeding of appropriately engineered bacterial strains may be used as described in Plaetinck et al., WO 00/01846.

Still another embodiment of the invention comprises incomplete rescue typically obtained by heterologous transgenes. For example, a strain *daf-16; daf-2; Ex[daf-16b::hsFKHR]* has been constructed in which *daf-16* loss of function, in itself rescuing from *daf-2* induced dauer formation, is rescued by the human homolog FKHR under the *C. elegans* *daf-16b* promoter. This rescue is incomplete, to about 60% dauer formation, so that 40% grow to adulthood

- 41 -

(Gary Ruvkun, personal communication). Any other homologue of *daf-16*, for example the human genes FKHL1 or AFX, or others, mammalian or human, could be used in combination of suitable promoters, either one
5 of the endogenous *daf-16* promoters, *daf-16a* or *daf-16b* or both, or a heterologous promoter, preferably with ubiquitous expression or nervous system expression.

Still another embodiment of the invention is based on the addition of pheromone preparations so
10 that the fraction of worms growing adults is driven below 60%, preferably below 40%, more preferably below 40%, such as between 10% and 20%. As already mentioned, Sze and co-workers (Nature 403: 560-564) generated a *tph-1(mg280)* mutation, which induces dauer
15 arrest at 15%, mimicking low food supply and with some resistance to temperature control. However, since the dauer arrest can be enhanced to 80% using a *daf-7* mutation, which are defective in production of a TGF β like molecule signalling the absence of pheromone,
20 addition of pheromone could achieve the desired level of 80% dauer formation as an alternative to the double mutant. Pheromone preparations may be obtained after the method of Golden & Riddle 1984, PNAS 81: 819-823.

This screening method of the invention is again
25 based on bypass of/release from the dauer larval state and there are several different ways of screening for bypass of/release from dauer which may be used in accordance with the invention, see below. The invention can as well be based on any other phenotype
30 relating to the insulin pathway, such as are observed in *daf-2* mutations, including but not exclusive to fat storage, regulation of metabolic enzymes or stress

- 42 -

resistance pathways or any other biochemically, transcriptionally or posttranscriptionally regulated effect that is measurable.

5 Set out below are ways of screening for bypass of or release from the dauer larval state which may be used in accordance with the invention.

 One of the simplest and most exact methods of, measuring bypass of/rescue from dauer larvae formation
10 is counting of adults. Counting of adults may be achieved using automated means, e.g. automatic plate readers, allowing the screen to be performed in mid-to-high throughput format in multiwell microtiter plates.

15 . A further method of screening for bypass of or rescue from the dauer phenotype exemplified herein is based on staining of adults using Nile Red and automated data acquisition (Example 2). Other methods of screening for release from the dauer larval state
20 are also encompassed by the invention.

 As an alternative to direct counting of adults indirect measurements, for example the consumption of food by measuring turbidity, may form a usable readout.

25

 Further methods of screening for bypass of/release from the dauer larval state are based on the use of reporter transgene. Suitable reporter transgene constructs generally comprise a promoter or
30 promoter fragment operably linked to a reporter gene.

 The promoter or promoter fragment is one which is capable of directing strong gene expression in adult

- 43 -

C. elegans but no or weak gene expression in dauer larvae, such as a promoter which is regulated by the *daf-2* signalling pathway (e.g. promoters regulated by the transcription factor *daf-16*) or vice versa (i.e. no or weak expression in adult, strong expression in dauer larvae. The term "operably linked" refers to a juxtaposition in which both components function in their intended manner, i.e. the promoter drives expression of the reporter gene. One example of a suitable transgene is a construct comprising the *C. elegans vit-2* promoter operably linked to a luciferase reporter gene. Any other promoter that shows strong expression in adults but no or weak expression in dauer larvae may be used as an alternative to the *vit-2* promoter. Other reporter genes may be used as alternatives to luciferase. Preferably the reporter gene will be one encoding a product which is directly or indirectly detectable in the worm, for example a fluorescent, luminescent or coloured product, e.g. GFP or lacZ. Preferably expression of the reporter gene product in the worm will be measurable using an automated plate reader.

The inventors provide methods for constructing *ctl-1::luciferase* and a *sod-3::luciferase* reporter transgenes, the *ctl-1* and *sod-3* genes encoding respective a cytosolic catalase with markedly increase expression in *daf-2* dauer larvae (Taub et al. 1999, Nature 399:162-166) and a manganese superoxide dismutase strongly up-regulated in *daf-2* mutant adults (Honda and Honda 1999, FASEB 13: 1385-1393). The regulation of a mitochondrial manganese superoxide dismutase by *daf-2* is of particular interest, since it

- 44 -

has recently been shown that overexpression of a Mn-SOD in vascular endothelial cells can suppress several pathways involved in hyperglycaemic damage, indicating that those damages are caused by production
5 of superoxides (Nishikawa et al. 2000, Nature 404: 787-790).

To perform a screen using a reporter transgene the transgene must first be introduced into the *C. elegans* used in the screen. This may be achieved
10 using standard techniques for the construction of transgenic *C. elegans* well known in the art and described, for example, in Methods in Cell Biology, Vol 48, Ed. H.F.Epstein and D.C.Shakes, Academic Press.

15

Table 1: targets of the insulin receptor pathway

Targets	Human homologs	Function	Validation	Desired intervention
DAF-2	IR	Receptor tyrosin kinase	e1391 equals het. mutation of an morbidly obese diabetic patient	+
	PTP-1B	Protein tyrosin phosphatase	Mouse k.o. insulin hypersensitive	B
DAF-2	IRS-1, -2	Insulin receptor substrate	IR/+; IRS-1/+ age onset diabetes, IRS2 diabetic	+
AGE-1	p110	PI3-kinase catalytic subunit	p110 β insulin responsive	+
	p85/p55	PI3-kinase regulatory subunit	p85 α k.o. insulin hypersensitive	+/B
DAF-18	PTEN	PI-3' phosphatase	maternal and zygotic minus rescues <i>daf-2(e1370)</i>	B
	SHIP2	PI-5' phosphatase	Overexpression inhibits AKT activation	B
PDK-1	PDK1	AKT phosphorylation	gf rescues dauers, lf induces dauers	+
AKT-1, AKT-2	AKT = PKB	Forkhead TF phosphorylation	gf rescues, double RNAi induce dauers	+
DAF-16	EKHR, FKHL1	Transkription factor	lf rescues <i>daf-2 (e1370)</i>	B

- 46 -

The present invention will be further understood with reference to the following Experimental examples, together with the accompanying Figures in which:

5 Figure 1 illustrates the insulin receptor signalling pathway of *C. elegans*.

Figure 2 is a print of the acedb database entry on *daf-2*.

10

Figure 3 is a graph to show that vanadates can rescue the genetic insulin resistance caused by *daf-2* mutations in *C. elegans* in an assay based on bypass of/rescue from the dauer larval state.

15

Figure 4 is a graph to show that wortmannin further enhances insulin resistance caused by *daf-2* mutations in *C. elegans* in an assay based on bypass of/rescue from the dauer larval state.

20

Figure 5 scatter plot of mean and variance of controls for the screening experiment described in Example 1 (a) screening, (b) DRC.

25

Figure 6 shows distribution of controls and a maximum likelihood of fit of a negative binomial distribution for data generated in the screening experiment described in Example 1.

30

- 47 -

Figure 7 shows distribution of controls in % of the average of the plate for data generated in the screening experiment described in Example 1.

5

Figure 8 shows the results of a representative Nile red staining experiment (Example 2).

Figure 9 is a representation of pGQ1.

10

Figure 10 is a representation of pDW2020.

Figure 11 shows the complete nucleotide sequence of pDW2020.

15

Figure 12 shows the complete nucleotide sequence of pGQ1.

Figure 13 is a print of the acedb database entry on *ctl-1*.

20

Figure 14 is a representation of pGQ2.

Figure 15 is a representation of pCluc6.

25

Figure 16 shows the complete nucleotide sequence of pCluc6.

Figure 17 shows the complete nucleotide sequence of pGQ2.

30

- 48 -

Figure 18 is a print of the acedb database entry on
sod-3.

Figure 19 is a representation of pGQ3.

5

Figure 20 shows the complete nucleotide sequence of
pGQ3.

Figure 21 is a representation of pGQ4.

10

Figure 22 shows the complete nucleotide sequence of
pGQ4.

Figure 23 illustrates the cloning of pCluc6.

15

**Example 1: screening 23,040 compounds for activity in
the insulin-receptor pathway.**

20 **Materials used**

- 9cm plates seeded with OP50,
- three weeks old stock plates of *daf-2(m41)*
- M9 buffer
- S-complete buffer
- 25 • 96-well plates flat bottom NUCLON Surface
- 96-well plates U-bottom for dilutions compounds
- HB101 bacteria (routinely available)
- compounds (80 per 96-well plates) 10mM concentration
in 100% DMSO

30

- 49 -

Method

Test of the batch of bacteria to be used as food:

- Growth of HB101
 - fill a 2 liter Erlenmeyer sterile with 0,5l DYT
5 medium
 - inoculate with *E-coli* HB101 single colony
 - let shake for 24 hours at 250 rpm and 37 C
 - centrifuge in sterile 250ml centrifuge tubes 10
min 10000rpm.
 - 10 - resuspend in 120 ml S-basal medium (pipette up
and down and shake)
 - transfer to 8 15ml falcon tubes that were weighed
in advance
 - centrifuge second time 10 min 6000rpm
 - 15 - weigh the pellet
 - store at 4 C
- Test of the batch:
 - chunk a couple of plates of *m41*
 - bleach plates after 4 days, let eggs hatch on
20 unseeded plate at 15 C
 - wash off L1's after one night
 - bring 50 L1 in 80 µl S-complete in one 96 well
plate
 - add 10 µl 2% DMSO
 - 25 - add 10µl of 1.25% of the batch of bacteria to be
tested
 - put plate in closed box in the 21 C incubator
 - check on number of dauers after three days of
growth, should be no more then .10
 - 30 - if the batch is approved, it can be stored
undiluted at 4 C for several weeks

- 50 -

Protocol

Thursday:

- chunk 9 cm plates (take 1 plate/96-well plate to be filled)
- 5 - grow in middle incubator at 15 C (preferably same shelf)

Monday : bleach plates

- wash off in M9
- 10 - 10 plates/falcon 15ml
- put washed off plates back in 15 C incubator (only uncontaminated ones)
- spin down at 1300rpm/3min
- suck off M9
- 15 - add bleach
- when most worms are broken, add sucrose, shake, add 2 ml M9
- spin at 1300rpm/3 min
- carefully remove eggs from bottom of layer of M9,
- 20 bring in new falcon
- add M9 to 15ml
- spin down 1300rpm/3min
- add M9
- spin down 1300rpm/3min
- 25 - suck away M9 to 1ml
- divide eggs from one falcon over 3 unseeded plates
- put plates at 15 C to let eggs hatch

30 Tuesday :

a) preparation of the compound-plates

- 51 -

- dilute aliquot of compound in 96-well plate to 200 μ M in S-buffer (DMSO conc. 2%).
- replicate plates: four plates 10 μ l 200 μ M compound per well
- 5 - write number and replicate number on plates
- if there was no DMSO in col 1 and 12 of the aliquoted plate it has to be added (add 11 μ l of 2% DMSO)
- write number of the plate and the replicate on
- 10 the lid of the plates

b) preparation of the worms solution

1) "bleached L1's"

- wash L1 off plates in S-complete, 4 plates/15ml
- 15 falcon
- spin down at 1300rpm/3min
- add fresh S-complete to 100ml
- count worms in 10 μ l
- keep worm suspension at 15 C while counting
- 20 - dilute further to approximately 50 worms/80 μ l, count again
- mix well

2) "washed L1's"

- 25 - wash off plates that were washed yesterday
- spin down (1300rpm/3min), add S-complete, wash twice
- filter suspension over 11 micron mesh over embroidery hoop into lid of 9cm plate
- 30 - wash L1's one more time
- dilute to 50 worms/80 μ l in the same way as bleached L1

- 52 -

c) Final steps:

- add 1.25% freshly diluted HB101 bacteria to worm suspension so that final concentration is 0.125%
5 (1 volume of bacteria to 8 of worms)
- add 90 µl of worm-bacteria suspension/well with electronic pipette
- put plates in closed boxes with wet tissues in 21°C incubator
- 10 - monitor temperature in control box in incubator while growing (try to put boxes at the same shelf, avoid contact of the boxes to metal of cooling device!)

15 Friday: Scoring:

1. count 8 negative control wells/plate
2. plot the average and variance of the negative controls from each plate
3. check for differences between boxes, differently
20 treated L1's and replicates
4. if necessary define several groups, remove outliers
5. make a distribution of the negative controls per group (plot # of wells to the number of
25 worms/well)
6. for each defined group: fit a negative binomial distribution to the negative controls and determine the number of adults for a cut-off confidentiality of about 1% and about 0.1% (both
30 sides for screen of dauer rescue and dauer enhancers)

- 53 -

7. screening for dauer rescue is possible if average of negative control is between 0 and 15 adults/well, screening for dauer enhancers is possible if the average is above 5
- 5 8. screen through the plates and count the wells with high number of adults
9. if the number of adults in the well is below the cut-off value leave it
- 10 10. if the number of adults is above or at the 1% cut-off value circle the well as positive (for each of the replicate with a different color) and write the number in the circle
11. if the number of adults is above the 0.1% cut-off value estimate the number of adults
- 15 12. Put the lids of the 4 replicates of the same plate on top of each other
13. Search for wells with 2 or more positives in the 4 (or 3) replicates
14. Write down the number of the adults of each of
20 the 4 (or 3) replicates

Robustness

While the controls active in the pathway show the sensitivity of the assay (see Figures 2 and 3), its
25 specificity is determined by testing a range of compounds outside the pathway. Together with the reference compounds acting in the insulin signalling pathway, of which only Wortmannin and vanadates were active, anti-diabetics with a mode of action outside
30 the insulin pathway, including 3 guanidine derivatives (acting on glucose uptake and metabolism), 5 PPAR γ ligands (stimulating adipocyte differentiation) and 6

- 54 -

sulphonylureas (which act by increasing insulin
 secretion) were tested. None was found to be active in
 the assay. Further confirmation of the specificity of
 the screen is derived from testing a library of 800
 5 compounds from Tocris-Cookson, containing mainly
 neurological actives, at 20 μ M in triplicates. Only 4
 compounds rescued dauer formation, a rate not higher
 than for random libraries (see results).

10

Table 2

Name of compound	supply	MW	drug class/ disease area/ action(s)	solvent	Concentrations tested in μ M- (lethal) rescue, <i>dauer enhancer</i>
Synthalin	ICN	354.5	guanidine derivative, also NMDA antagonist	DMSO	(333; 166.7; 83.3; 33.3); 20; 16.6; 8.3; 3.3
Metformin HCl (1,1- dimethylbiguanide)	Sigma	165.6	guanidine derivative, biguanides, MOA?: decrease hepatic glucose production	DMSO	333; 166.7; 83.3; 33.3; 20
Phenformin HCl (phenethylbiguanide)	Sigma	241.7	guanidine derivative, biguanides, MOA?: decrease hepatic glucose production	DMSO	333; 166.7; 83.3; 33.3; 20
HNMPA(AM)3	Calbioc hem	454.4	insulin receptor tyrosine kinase inhibitor	DMSO	20
Rapamycin	ICN	914.2	insulin signalling enhancer, inhibitor of the mammalian target of rapamycin (mTOR) which is a downstream target of Akt and implicated in Akt's negative regulation of insulin signalling i.e.	DMSO	33.3; 16.6; 8.3;

			serine/threonine phosphorylation of IRS-1		
Quercetin	Sigma	338.3	insulin signalling inhibitor, inhibitor of phosphatidylinositol 3-kinase and of several other ATP-requiring enzymes e.g. PI4K, PKC, EGFR, calcium, SERCA activator by interacting with nucleotide binding site to mask PLB inhibition	DMSO	20
okadaic acid	Calbioc hem	805	insulin signalling inhibitor, inhibits PP2A and PP1	DMSO	10; 5; 2.5; 0.6
PD 98059	Calbioc hem	267.3	insulin signalling inhibitor, MEK1 inhibitor	DMSO	20
Wortmannin	Sigma	428.4	insulin signalling inhibitor, phosphatidylinositol 3-kinase inhibitor (potent and specific), inhibitor of neutrophil activation and of FMLP-mediated phospholipase D activation	DMSO	20
LY 294002	Sigma	307.3	insulin signalling inhibitor, phosphatidylinositol 3-kinase inhibitor (specific)	DMSO	100, 20
phorbol 12-myristate 13-acetate (PMA)	Biomol	616.8	insulin signalling inhibitor, PKC activator (elicits serine/threonine phosphorylation of IRS-1)	DMSO	20
Phosphatidylinositol-3,4,5-trisphosphate [stearyl, arachidonoyl, tetraammonium salt)	Alexis	1123.1	insulin signalling, identical to endogenous PI(3,4,5)P3 (not an analog containing only saturated fatty acid residues, therefore greater biological activity), activates Ca ²⁺ -insensitive PKC, activates Akt (a serine/threonine kinase) by directly interacting with the Akt pleckstrin homology (PH) domain	DMSO	2.8; 1.4; 0.7

Phosphatidylinositol-3,4-bisphosphate [L-alpha-] (dipalmitoyl, pentaammonium salt)	Calbioc hem	1056.2	insulin signalling, mimics endogenous PI(3,4)P2, activates Ca2+-insensitive PKC, activates Akt (a serine/threonine kinase) by directly interacting with the Akt pleckstrin homology (PH) domain	DMSO	3.17; 1.9; 1.58; 0.79
Phosphatidylinositol-3,4,5-trisphosphate [L-alpha-] (dipalmitoyl, heptaammonium salt)	Calbioc hem	1170.2	insulin signalling, mimics endogenous PI(3,4,5)P3, activates Ca2+-insensitive PKC, activates Akt (a serine/threonine kinase) by directly interacting with the Akt pleckstrin homology (PH) domain	DMSO	2.96; 1.74; 1.48
Thalidomide	ICN	258.2	insulin signalling, TNF inhibitor	DMSO	333; 166.7; 83.3; 33.3; 20
Perhexiline	Sigma	393.6	insulin, carbohydrate metabolism, inhibitor of myocardial carnitine palmitoyltransferase-1 ("antidiabetics"), sodium, calcium, dual Na+/Ca2+ (T-type) channel blocker, anti-angina (coronary vasodilator), diuretic	DMSO	(333; 166.7; 83.3; 33.3); 20; 16.6; 8.3; 3.3
L-arginine	Sigma	174.2	nitric oxide, insulin secretagogue (NO dependent)	water	333; 166.7; 83.3; 33.3; 20
D-arginine	Sigma	174.2	nitric oxide, negative control of L-arginine (insulin secretagogue)	water	20
LY 171883	Sigma	318.4	PPARgamma activator (weak), selective LTD4 antagonist	DMSO	20
linoleic acid (9,12-octadecadienoic acid)	Sigma	280.4	PPARgamma ligand	DMSO	(333; 166.7; 83.3; 33.3); 20; 16.6; 8.3; 3.3
Linolenic acid (9,12,15-octadecatrienoic acid)	Sigma	278.4	PPARgamma ligand	DMSO	(333; 166.7; 83.3; 33.3); 20; 16.6; 8.3; 3.3
Eicosatetraynoic acid [5,8,11,14-] (ETYA)	ICN	296.5	PPARgamma ligand, insulin sensitizers, eicosanoid	DMSO	333; 166.7; 83.3; 33.3; 20

Rosiglitazone (BRL49653)		359	PPARgamma-specific agonist (insulin-sensitizing properties, used in type II diabetes)	water	909; 500; 263; 135; 55; 27.6; 13.85
Chelerythrine chloride	Sigma	383.8	protein kinase C inhibitor (potent, selective, IC ₅₀ 0.7 μM)	DMSO	10
Cantharidic acid	Sigma	214.2	protein phosphatase 2A inhibitor (IC ₅₀ 53 nM)	DMSO	20
Phenylarsine oxide	Calbiochem	168	PTP inhibitor, also inhibits PI3-kinase activity	DMSO	20
Bromotetramisole oxalate [L-p-]	Biomol	373.2	PTP inhibitor, also well known inhibitor of alkaline phosphatase, mimics the action of orthovanadate in the potentiation of fluorouracil antiproliferative activity	water	20
Bromotetramisole oxalate [D-p-]	Biomol	373.2	PTP inhibitor, also well known inhibitor of alkaline phosphatase, mimics the action of orthovanadate in the potentiation of fluorouracil antiproliferative activity; inactive isomer, negative control	water	20
Dephostatin	Calbiochem	168.2	PTP inhibitor, IC ₅₀ 7.7 μM, also nitric oxide donor (stable NO donor for S-nitrosation of proteins)	DMSO	333; 166.7; 83.3; 20
vanadium(II) chloride	Aldrich-Sigma	121.85	PTP inhibitor, vanadium compound	DMSO	20
vanadium(III) chloride	Aldrich-Sigma	157.3	PTP inhibitor, vanadium compound	DMSO	1000; 500; 250; 100; 20
vanadium(III) oxide	Aldrich-Sigma	149.88	PTP inhibitor, vanadium compound	DMSO	20
vanadium(IV) oxide	Aldrich-	165.88	PTP inhibitor, vanadium compound	DMSO	20

	Sigma				
vanadium(V) oxide	Aldrich-Sigma	181.88	PTP inhibitor, vanadium compound	DMSO	20
vanadyl sulfate	Aldrich-Sigma	163	PTP inhibitor, vanadium compound	DMSO	1000; 500; 250; 100; 20
vanadyl trifluoride	Fluka-Sigma	123.94	PTP inhibitor, vanadium compound	DMSO	20
mpV (Pic) (mono peroxy (picolinato) oxovanadate(V))	Calbioc hem	257.1	PTP inhibitor, vanadium compound	DMSO	1000; 500; 250; 100; 20
sodium metavanadate	Sigma	121.9	PTP inhibitor, vanadium compound, also inhibits ATPase and alkaline phosphatase	water	1000; 500; 250; 100; 20
sodium orthovanadate	Sigma	183.9	PTP inhibitor, vanadium compound, also inhibits ATPase and alkaline phosphatase	water	1000; 500; 250; 100; 20
bpV (Phen) (Potassium Bisperoxy (1,10-phen anthroline) oxovanadate(V))	Calbioc hem	404.3	PTP inhibitor, vanadium compound, potent	DMSO	1000; 500; 250; 100; 20
bpV(bipy) (potassium bisperoxy(bipyridine) oxovanadate(V))	Alexis	326.2	PTP inhibitor, vanadium compound, potent	DMSO	1000; 500; 250; 100; 20
bpV(Hopic) (di potassium bis peroxy(5-hydroxy pyridine-2-carboxyl)-oxovanadate(V))	Alexis	347.2	PTP inhibitor, vanadium compound, potent	DMSO	1000; 500; 250; 100; 20
bpV(pic)	Alexis	367.3	PTP inhibitor, vanadium compound,	DMSO	1000; 500; 250;

(dipotassium bisperoxo(picolinato)oxovanadate(V))			potent		100; 20
acetohexamide	ICN	324.4	sulfonylureas, first generation, MOA: insulin secretagogue by blocking K ⁺ (ATP) channels	DMSO	333; 166.7; 83.3; 33.3; 20
chlorpropamide	Sigma	276.7	sulfonylureas, first generation, MOA: insulin secretagogue by blocking K ⁺ (ATP) channels	DMSO	333; 166.7; 83.3; 33.3; 20
tolazamide	Sigma	311.4	sulfonylureas, first generation, MOA: insulin secretagogue by blocking K ⁺ (ATP) channels	DMSO	333; 166.7; 83.3; 33.3; 20
tolbutamide	Sigma	270.3	sulfonylureas, first generation, MOA: insulin secretagogue by blocking K ⁺ (ATP) channels	DMSO	333; 166.7; 83.3; 33.3; 20
glipizide	RBI	445.53	sulfonylureas, second generation, MOA: insulin secretagogue by blocking K ⁺ (ATP) channels	DMSO	333; 166.7; 83.3; 33.3; 20
glyburide (glybenclamide)	Tocris	494.1	sulfonylureas, second generation, MOA: insulin secretagogue by blocking K ⁺ (ATP) channels	DMSO	333; 166.7; 83.3; 33.3; 20
diazoxide	Tocris	230.7	potassium, K ⁺ channel opener, activates ATP-sensitive K ⁺ channels, antihypertensive, also stimulates K ⁺ channels in pancreatic islet cells (prodiabetic side effects), diabetes	DMSO	333; 166.7; 83.3; 33.3; 20

- 60 -

Data acquisition

All screening was done at 20 μ M compound concentration in quadruplicates, except 2000 compounds of Diverset in triplicates. Confirmation was done at 4
5 concentrations. Questionable dose responses were repeated, if necessary at lower concentrations and/or 2 fold dilution steps. All worms that bypassed dauer stage, L4s and adults, were counted under a Leica MZ12 dissection scope and together referred to as number of
10 adults per well. First, the 8 negative controls (column 1) of all plates were counted, typically between 800 and 1280 (25 to 40 plates times 4 per screening session). Data were transferred to Excel files and average and variance of the 8 controls of
15 each plate calculated and plotted.

Outliers of unusual high average or variance were removed for calculation, since they were found to have an inappropriately large effect on the calculations
20 below (3 plates in the example of Figure 5a). Counting errors were found to have a rather weak effect. The number of wells was plotted against the number of adults per well and a negative binomial distribution fitted by maximum likelihood. In some cases it was
25 necessary to split a session in two or three different subsessions mainly due to differences in incubator location or worm handling.

Then the number of adults per well where the
30 cumulative negative binomial distribution was closest to 99% was determined and referred to as 1% cut-off. In the example shown in Figure 6, 20 adults per well

- 61 -

were at 1.10% indicating that the probability to have 20 or more adults per well is 1.10%. This calculates to a 4% chance for a single false positive in quadruplicates, but only to a 0.07% chance for a double false positive. Therefore a compound is positive, if at least 2 replicates have values at the cut-off or higher. In addition the 0.1% cut-off was determined similarly (24 adults in the example shown in Figure 6) and if at least 2 replicates were reaching that stronger value the compound was referred to as strong positive.

The plates were then screened through quickly to find wells with a high number adults, which were counted and if found to reach the cut-off value the position on the lid was circled and the exact value written in the circle. For higher numbers above the 0.1% cut-off an estimate rather than an exact count proved sufficient. Finally the transparent lids of the 4 replicate plates were stacked on top of each other and by looking through them it was determined whether 2 or more lids were circled in any position. For those positions all the positive values were written into an excel file.

For confirmation by dose response fresh compound in 100% DMSO was used and from an initial dilution to 2% DMSO three further dilutions in 3.16 fold steps with a 2% DMSO solution in S-buffer were prepared. In that way 4 concentrations, 20 μ M, 6.3 μ M, 2 μ M and 0.63 μ M were tested, all in 0.2% DMSO background. Both columns 1 and 12 contained 0.2% DMSO as control. Each plate

- 62 -

contained 20 different compounds, with 4 replica-plates of them.

Table 3

	1	2	3	4	5	6	7	8	9	10	11	12
	comp1	comp2	comp3	Comp4	comp5	comp6	comp7	comp8	comp9	comp1		
A	cntrl	20µM	20µM	20µM	20µM	20µM	20µM	20µM	20µM	20µM	20µM	cntrl
B	cntrl	6µM	6µM	6µM	6µM	6µM	6µM	6µM	6µM	6µM	6µM	cntrl
C	cntrl	2µM	2µM	2µM	2µM	2µM	2µM	2µM	2µM	2µM	2µM	cntrl
D	cntrl	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	cntrl
E	cntrl	20µM	20µM	20µM	20µM	20µM	20µM	20µM	20µM	20µM	20µM	cntrl
F	cntrl	6µM	6µM	6µM	6µM	6µM	6µM	6µM	6µM	6µM	6µM	cntrl
G	cntrl	2µM	2µM	2µM	2µM	2µM	2µM	2µM	2µM	2µM	2µM	cntrl
H	cntrl	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	0.6µM	cntrl
	comp1	comp1	comp1	Comp1	comp1	comp1	comp1	comp1	comp1	comp1	comp2	
	1	2	3	4	5	6	7	8	9	10	11	12

5

"Cntrl"-abbreviation for control

For some compounds an additional dose response with 7 concentrations was made, mostly with 2 fold dilutions to obtain 20 µM, 10 µM, 5 µM, 2.5 µM, 1.25 µM, 0.63 µM and 0.31 µM. In that case also row H contained controls. Each plate contained 10 different compounds, with 4 replica-plates of them. An example of the 26

- 63 -

negative controls of 16 plates shows the variability of the mean while the standard deviation remained fairly constant (Figure 5b). Furthermore, the negative controls expressed as percentage of the plate mean were approximately normal distributed (Figure 7). Therefore all data were normalized according to the calculation below, which centers value of no effect at 0 and calibrates the y-axis to standard deviations. The concentrations are on the x-axis in logarithmic scale. All 4 replicates are plotted, in addition a smoothed line through the averages is plotted.

value in SD = (number of adults of the well -1)/SD of the controls of the set
average controls of the plate

15

A compound was determined as confirmed and designated a hit when either the average or two of the 4 values reached 2.5 SD (corresponds to 99.3% confidence) at any concentration and a reasonable dose-response is apparent.

20

Results

From 23.040 compounds a total of 300 positives were obtained during the screening, of which 173 could be reconfirmed.

25

- 64 -

Table 4

library name	size	Positives	confirmed hits	% re- confirmed	hit rate
Library 1	2000	33	3	9%	0.15%
Library 2	5040	92	62	67%	1.23%
Library 3	16000	175	108	62%	0.68%
TOTAL	23040	300	173	57%	0.75%

To estimate the potency of the screen, that is to
5 estimate what fraction of compounds that could have
been identified with the assay have actually been
identified during the screen, an analysis on 47
compounds defining 11 chemical clusters has been
performed: 36 of these compounds have been confirmed.
10 Another 40 compounds, which were not found to be
active in the original screen but are members of those
clusters, were submitted to dose response
confirmation. 4 more hits have been identified. In
total 40 compounds could be confirmed, 36 of the
15 screen positives and 4 from the extra set. Hence 90%
of the final hits of these clusters were detected in
the original screen and 10% were missed.

- 65 -

Table 5

Cluster	positives	confirmed hits	similar negatives	extra hits	final hits
1	5	4	1	0	4
3	6	6	7	1	7
4	7	6	1	0	6
5	4	4	1	0	4
6	3	3	5	1	4
7	5	3	1	0	3
8	3	1	7	1	2
9	5	4	13	0	4
12	5	2	1	0	2
13	2	2	2	0	2
15	2	1	1	1	2
Total	47	36	40	4	40

Conclusions

1. A mutation in the *C. elegans* insulin receptor, *daf-2(m41)*, was used successfully in an pharmacological assay for compounds acting in the downstream pathway.
2. The assay is sensitive enough to screen at 20 μ M compound concentrations, at which there were

- 66 -

nearly no problems due to lethality (27 of 23,040).

3. A hit rate of 0.75% from combinatorial chemistry libraries has been obtained, strongly dependent
5 on the library.
4. The screen is specific for the insulin receptor pathway and is unlikely to yield many hits upstream e.g. stimulating insulin release.
5. The active compounds are candidates to cure
10 insulin resistance and therefore of potential therapeutic use in type II diabetes and obesity.
6. Since the compounds bypass the need of insulin they are also of potential use in type I diabetes.
- 15 7. The major mode of compound entry in *C. elegans* is the gut which pre-selects for orally active compounds.
8. The activity is retrieved from a whole-organism readout leaving intact tissue-specific insulin
20 signalling and feedback loops.

- 67 -

Table 6: Retest of 94 compounds at 20µM on 3 different daf-2 alleles, m41 at 211C, e1368 and e1370 at 251C.

Values: 3: all replicates above 99% threshold, 2: median above 99.9% threshold, 1: median above 99% threshold, 0: median below 99% threshold.

ID	MW	Plat	Row	Col	m41	e1368	e1370
e							
217485	547.18	1	A	2	1	1	0
211706	472.55	1	A	3	3	3	0
181141	459.51	1	A	4	3	1	0
259910	384.53	1	A	5	0	0	0
194326	393.49	1	A	6	2	0	0
217336	420.04	1	A	7	3	3	0
267546	372.51	1	A	8	0	0	0
228433	405.56	1	A	9	0	0	0
264792	436.94	1	A	10	3	0	0
255126	431.50	1	A	11	3	0	0
100718	399.88	1	B	2	3	0	0
182576	486.39	1	B	3	0	0	0
232839	475.30	1	B	4	3	1	0
217339	394.00	1	B	5	3	1	0
217341	394.00	1	B	6	3	2	0
118776	437.52	1	B	7	2	0	0
118783	452.35	1	B	8	3	2	0
118789	442.35	1	B	9	2	1	0
248144	440.89	1	B	10	3	0	0
234291	462.76	1	B	11	0	0	0
212465	367.39	1	C	2	0	0	0
144331	363.98	1	C	3	0	0	0
138263	372.51	1	C	4	2	1	0
264982	352.48	1	C	5	1	1	0
267659	386.93	1	C	6	1	0	0
115771	391.50	1	C	7	3	0	0
105359	326.40	1	C	8	3	0	0
267467	419.37	1	C	9	0	0	0
236867	480.25	1	C	10	0	0	0
225671	365.44	1	C	11	0	0	0
225858	444.33	1	D	2	0	1	0
225615	523.23	1	D	3	0	1	0
101025	431.42	1	D	4	1	0	0
255192	420.38	1	D	5	3	1	0
217850	391.27	1	D	6	3	0	0
214475	329.36	1	D	7	3	1	0
114446	479.71	1	D	8	2	0	0
261736	378.40	1	D	9	2	0	0
210145	373.84	1	D	10	0	0	0
114816	304.40	1	D	11	2	0	0

- 68 -

210877	445.34	1	E	2	0	0	0
189119	379.38	1	E	3	3	1	0
203845	379.38	1	E	4	1	0	0
190303	303.36	1	E	5	0	0	0
253121	524.23	1	E	6	3	1	0
228525	462.45	1	E	7	2	1	0
118761	381.89	1	E	8	2	0	0
228489	428.55	1	E	9	1	0	0
250480	332.36	1	E	10	2	1	0
118765	416.33	1	E	11	3	0	0
254230	425.24	1	F	2	0	0	0
255339	427.69	1	F	3	2	1	0
250001	383.24	1	F	4	2	0	0
255335	383.24	1	F	5	2	2	0
263986	330.86	1	F	6	0	0	0
236861	486.21	1	F	7	0	0	0
104926	280.35	1	F	8	0	1	0
133891	272.30	1	F	9	0	0	0
154290	364.27	1	F	10	2	0	0
189005	363.76	1	F	11	1	0	0
195094	346.29	1	G	2	2	0	0
203897	408.21	1	G	3	3	0	0
210775	510.21	1	G	4	1	0	0
214387	376.64	1	G	5	3	0	0
219414	318.33	1	G	6	1	0	0
228301	311.36	1	G	7	0	0	0
228488	414.53	1	G	8	1	0	0
230672	376.21	1	G	9	0	0	0
231561	365.88	1	G	10	0	0	0
236341	386.41	1	G	11	0	0	0
249726	422.19	1	H	2	1	0	0
249746	373.33	1	H	3	2	0	0
253051	311.57	1	H	4	0	0	0
257516	380.73	1	H	5	0	0	0
258687	305.36	1	H	6	0	0	0
260067	357.18	1	H	7	0	0	0
265080	346.29	1	H	8	0	1	0
268434	372.42	1	H	9	0	0	0
273546	443.05	1	H	10	0	0	0
276545	337.70	1	H	11	1	0	0
278617	430.05	2	A	2	0	0	0
279528	316.34	2	A	3	0	0	0
281078	344.25	2	A	4	3	0	0
283400	390.31	2	A	5	0	0	0
284204	301.26	2	A	6	0	0	0
284316	385.22	2	A	7	0	0	0
286676	354.15	2	A	8	0	0	0
301158	475.86	2	A	9	3	2	0
304896	432.26	2	A	10	0	0	0
307069	362.82	2	A	11	0	0	0
309471	453.32	2	B	2	0	0	0
310513	318.13	2	B	3	2	1	0
313944	416.29	2	B	4	0	0	0

- 69 -

316982	516.85	2	B	5	2	0	0
number or compounds active					53	25	0
percentage of compounds active					56%	27%	0%

5 Example 2: automatic data aquisition with Nile Red
 staining

Material:

10 Hardware:

- ```
- microtiterplates:96 well black U-shaped plates
 (DYNEX Microfluor7 2)
- Wallac 1420 plate reader (Victor 2):
 Nile Red protocol: excitation = 530 nm
 emission = 590 nm
```

15 emission = 590 nm

```
. Counting time: 1 second
```

CW lamp energy: 30445

Emission aperture: damp

Counter position: top

20            Measurement height: 3 mm from bottom of the plate

Consumables:

- Nile Red (Sigma, N-3013)
- Ivermectin (ICN, 196009)

25

- 70 -

**Method:**

- Prepare a 100 mM solution of Nile Red (Nile Blue A Oxazone) in pure methanol. Centrifugate to remove the saturated solution from the  
5 undissolved Nile Red.
- Dilute in steps of 10 with buffer to 500  $\mu$ M.
- Add 1:1 Nile Red to the worms and incubate for 30 min at room temperature.
- Add 10  $\mu$ M ivermectin final concentration and  
10 incubate for 30 min at room temperature.
- Measure.

15 **Example 3: automatic data aquisition with a vit-2::luciferase reporter**

**Material:**

## Hardware:

- microtiterplates: 96 well white U-shaped plates  
20 (DYNEX Microfluor  $\hat{a}$  2)
- Wallac 1420 plate reader (Victor 2):  
Luciferase protocol  
Emission Filter: no filter  
Counting time: 3 seconds  
25 Emission aperture: normal

**Consumables:**

- Triton X-100 (BDH, 306324N)
- Dual-Luciferase  $\hat{a}$  Reporter Assay System (Promega,  
30 E4550)

- 71 -

**Method:**

- Add Triton X-100 (1% final concentration) to lyse the worms.
- Shake for 1 minute and freeze.
- 5 - Thaw the plates and add 1:1 luciferine.
- Shake for 1 minute and measure.

**Example 4: construction of ctl-1::luciferase and**  
10 **sod-3::luciferase reporters**

## 1) Construction of pGQ1

## 1.1 PCR

15

PCR (turbo pfu) on N2 genomic DNA with:

oGQ1:ctl-1::GFP fw (PstI):

5' AAAACTGCAGCCAATGCATTGGAAGAGATATTTGCGCGTCAAATATGTTTGTGTCC3'

oGQ2bis:ctl-1::GFP rv (BamHI)

20

5'CGCGGATCCGGCCGATTCTCCAGCGACCG3'

## 1.2 Cloning

- Digest of the PCR fragment with PstI and BamHI

- Ligation into pDW2020 and transformation into DH10B

25

## 2) Construction of pGQ2

## 2.1 PCR

30

PCR (turbo pfu) on N2 genomic DNA with:

oGQ3:ctl-1::luciferase fw (StuI):

5' CCAGGCCTGAGATATTTTGC GCGTCAAATATGTTTGTGTCC3'

oGQ4:ctl-1::luciferase rv (SacI)

- 72 -

5'CGGAGCTCCGATTGGATGTGGTGAGCAGG3'

## 2.2 Cloning

- Digest of the PCR fragment with StuI and SacI
- 5 - Ligation into pCluc6 and transformation into DH10B

## 3) Construction of pGQ3

### 10 3.1 PCR

PCR (turbo pfu) on N2 genomic DNA with:

oGQ7:sod-3 fw:

5'GCAGAATTTGCAAAACGAGCAGGAAAGTC3'

oGQ6:sod-3::luciferase rv (AscI)

15 5'TTGGCGCGCCAAGCCTTAATAGTGTCCATCAGC3'

### 3.2 Cloning

- Digest of the PCR fragment with PstI and AscI
  - Ligation into pDW2020 and transformation into HD10B
- 20

## 4) Construction of pGQ4

### 4.1 PCR

25

PCR (turbo pfu) on N2 genomic DNA with:

oGQ7:sod-3 fw:

5'GCAGAATTTGCAAAACGAGCAGGAAAGTC3'

oGQ8:sod-3::luciferase rv (SacI)

30 5'CTGAGCTCGGCTTAATAGTGTCCATCAGC3'

### 4.2 Cloning

- Digest of the PCR fragment with PstI and SacII

- 73 -

- Ligation into pCluc6 and transformation into HD10B

#### Example 5: Construction of pCluc6

5 Vector:

- Restriction digest of pCluc2 with HindIII
- Purification, protocol: Jetsorb

Insert:

- PCR the vit-2 promoter (248 bp in front of exon1  
10 just before ATG ) with primers (designed from ACeDB  
C42D8.2) that contain HindIII RE sites out of N2  
genomic DNA:

vit-2F: 5'CCCCCAAGCTTCCATGTGCTAGCTGAGTTTCATCATGTCC3'

vit-2R: 5'CCCCCAAGCTTGGCTGAACCGTGATTGG3'

- 15 - Restriction digest on PCR product with HindIII
- Purification, protocol: Jetsorb

pCluc6:

- T4 DNA ligation of vector and insert
- 20 - Transformation into DH10B
- Mini DNA preparation, protocol: Wizard SV Miniprep
- determine direction of insert by RE cleavage  
XbaI/NheI
- Maxi DNA preparation, protocol: Jetstar
- 25 - Check maxiprep by sequencing with o-PUCI primer.

#### Standard methods and worm strains

- Standard methods for culturing nematodes are described
- 30 in Methods in Cell biology Vol. 48, 1995, ed. by  
Epstein and Shakes, Academic press. Standard methods  
are known for creating mutant worms with mutations in

- 74 -

selected *C. elegans* genes, for example see J. Sutton  
and J. Hodgkin in "The Nematode *Caenorhabditis*  
*elegans*", Ed. by William B. Wood and the Community of  
*C. elegans* Researchers CSHL, 1988 594-595; Zwaal et  
5 al, "Target - Selected Gene Inactivation in  
*Caenorhabditis elegans* by using a Frozen Transposon  
Insertion Mutant Bank" 1993, Proc. Natl. Acad. Sci.  
USA 90 pp 7431 -7435; Fire et al, Potent and Specific  
Genetic Interference by Double-Stranded RNA in *C.*  
10 *elegans* 1998, Nature 391, 860-811. A population of  
worms can be subjected to random mutagenesis by using  
EMS, TMP-UV or radiation (Methods in Cell Biology, Vol  
48, *ibid*). Several selection rounds of PCR could then  
be performed to select a mutant worm with a deletion  
15 in a desired gene.

A range of specific *C. elegans* mutants are available  
from the *C. elegans* mutant collection at the *C.*  
*elegans* Genetic Center, University of Minnesota, St  
20 Paul, Minnesota.

*E. coli* strain OP50 can be obtained from the *C.*  
*elegans* Genetics Center, University of Minnesota, St  
Paul, Minnesota, USA.

25

- 75 -

**CLAIMS:**

1. A method for the identification of a compound which is capable of modulating insulin signalling pathways, which method comprises:  
5 providing *C. elegans* dauer larvae;  
contacting said larvae with a test compound; and  
screening for release from the dauer larval state, wherein the *C. elegans* dauer larvae possess a sensitized genetic background, as compared to the  
10 reference *daf-2* mutant *el370*.
2. Method according to claim 1, in which the dauer larvae belong to a nematode strain which has an  
15 Insulin Sensitivity Value ("ISV") that is greater than the ISV for the reference nematode strain CB1370, in particular more than 1% greater, preferably more than 5% greater, more preferably more than 10% greater, even more preferably more than 20% greater.  
20
3. Method according to claim 1 and/or 2, in which the dauer larvae belong to a nematode strain which has an ISV that is >30 %, preferably >40%, even more preferably >50%.  
25
4. A method as claimed in claim 1 wherein the *C.elegans* dauer larvae are *daf-2(m41)* mutants.
5. A method as claimed in claim 1 wherein the  
30 *C. elegans* dauer larvae comprise a *daf-2* class I allele other than *daf-2(m41)*.

- 76 -

6. A method as claimed in claim 1 wherein the  
*C. elegans* dauer larvae comprise at least one loss-of-  
function or reduction-of-function mutation in a  
5 gene(s) downstream of the insulin receptor in the  
insulin signalling pathway.

7. A method as claimed in claim 6 wherein the  
*C. elegans* dauer larvae comprise a loss-of-function or  
10 reduction-of-function mutation in the *age-1* gene.

8. A method as claimed in claim 6 wherein the  
*C. elegans* dauer larvae comprise loss-of-function or  
reduction-of-function mutations in the *akt-1* gene and  
15 the *akt-2* gene.

9. A method as claimed in claim 6 wherein the  
*C. elegans* dauer larvae comprise a loss-of-function or  
reduction-of-function mutation in the *pdk-1* gene.  
20

10. A method as claimed in claim 9 wherein the  
*C. elegans* dauer larvae are *pdk-1(sa680)* mutants.

11. A method as claimed in claim 1 wherein the  
25 *C. elegans* dauer larvae are larvae wherein the dauer  
phenotype is induced by treatment with an inhibitor  
inhibitor of at least one component of the insulin  
receptor signalling pathway.

30 12. A method as claimed in claim 11 wherein the  
inhibitor compound is an inhibitor of the *C. elegans*  
PI3-kinase.

- 77 -

13. A method as claimed in claim 12 wherein the inhibitor compound is wortmannin or LY294002.

5           14. A method as claimed in claim 1 wherein expression of at least one gene downstream of the insulin receptor in the insulin receptor signalling pathway in said *C. elegans* dauer larvae is inhibited by RNAi inhibition.

10

          15. A method as claimed in claim 1 wherein the *C. elegans* dauer larvae comprise a gain-of-function mutation in the *daf-16* gene.

15

          16. A method as claimed in claim 1 wherein the *C. elegans* dauer larvae comprise a gain-of-function mutation in the *daf-18* gene.

20

          17. A method as claimed in claim 1 wherein the *C. elegans* dauer larvae comprise a gain-of-function mutation in the *C. elegans* homologue of the SHIP2 gene.

25

          18. A method as claimed in claim 1 wherein the *C. elegans* larvae dauer comprise a gain-of-function mutation in the *C. elegans* homologue of the PTP-1B gene.

30

          19. A method as claimed in claim 1 wherein the *C. elegans* dauer larvae exhibit a defect in perception of environmental signals.

- 78 -

20. A method as claimed in claim 19 wherein the said *C. elegans* dauer larvae comprise a mutation in the *tph-1* gene.

5        21. A method as claimed in claim 20 wherein the said *C. elegans* dauer larvae are *tph-1(mg280)* mutants.

22. A method as claimed in claim 1 wherein the *C. elegans* dauer larvae comprise a *daf-c* mutation in a  
10 *daf* gene selected from the group consisting of *daf-1*, *daf-4*, *daf-7*, *daf-8*, *daf-11*, *daf-14*, *daf-21*, *daf-19* and *daf-28*.

23. A method as claimed in claim 1 wherein the  
15 *C. elegans* dauer larvae comprise a mutation in a gene encoding a neuronal G-protein.

24. A method as claimed in claim 1 wherein the  
20 *C. elegans* dauer larvae are *unc-64(e264)*; *unc-31* (*e928*) mutants.

25. A method as claimed in any one of claims 1 to 24 wherein the step of screening for release from the dauer larval state comprises screening for adult  
25 *C. elegans*.

26. A method as claimed in any one of claims 1 to 24 wherein the step of screening for release from the dauer larval state comprises screening for changes  
30 in fat storage.

- 79 -

27. A method as claimed in any one of claims 1 to 24 wherein said *C. elegans* dauer larvae further comprise a reporter transgene comprising a promoter which is capable of directing strong gene expression in adult *C. elegans* and no or weak expression in dauer larvae or vice versa operably linked to a reporter gene and the step of screening for release from the dauer larval state comprises screening for changes in expression of the said reporter gene.

10

28. A method for the identification of a compound which is capable of modulating insulin signalling pathways, which method comprises: providing *C. elegans* dauer larvae; contacting said larvae with a test compound; and screening for release from the dauer larval state, wherein conditions of the assay are selected such that a basal level of release from the dauer larval state is observed in the absence of the test compound.

20

29. A method as claimed in claim 28 wherein the basal level of release from the dauer larval state is between 0.1% and 40%.

30. A method as claimed in claim 29 wherein the basal level of release from the dauer larval state is between 1% and 30%.

31. A method as claimed in claim 30 wherein the basal level of release from the dauer larval state is between 2% and 20%.

30

- 80 -

32. A method as claimed in any one of claims 28 to 31 wherein the *C. elegans* dauer larvae are *daf-2(m41)* mutants.

5        33. A method as claimed in any one of claims 28 to 31 wherein the *C. elegans* dauer larvae are *daf-2; daf-18* double mutants.

10       34. A method as claimed in any one of claims 28 to 31 wherein the *C. elegans* dauer larvae are *Daf-d* mutants.

15       35. A method as claimed in any one of claims 28 to 31 wherein the *C. elegans* dauer larvae comprise a gain-of-function mutation in the *pdk-1* gene.

36. A method as claimed in claim 35 wherein the *C. elegans* dauer larvae are *pdk-1(mg142)* mutants.

20       37. A method as claimed in any one of claims 28 to 31 wherein the *C. elegans* dauer larvae comprise a gain-of-function mutation in the *akt-1* gene.

25       38. A method as claimed in claim 37 wherein the *C. elegans* dauer larvae are *akt-1(mg144)* mutants.

30       39. A method as claimed in any one of claims 28 to 31 wherein the *C. elegans* dauer larvae are *daf-16; daf-2* double mutants and further comprise a transgene capable of expressing a mammalian homolog of the *daf-16* protein.

- 81 -

40. A method as claimed in claim 39 wherein the mammalian homolog of the daf-16 protein is the human FKHR protein, the human FKHRL1 protein or the human AFX protein.

5

41. A method as claimed in claim 28 wherein said *C. elegans* dauer larvae are larvae which have been treated with pheromone to reduce that fraction of worms growing to adults to below 40%.

10

42. A method as claimed in claim 41 wherein said *C. elegans* dauer larvae are larvae which have been treated with pheromone to reduce that fraction of worms growing to adults to below 30%.

15

43. A method as claimed in claim 42 wherein said *C. elegans* dauer larvae are larvae which have been treated with pheromone to reduce that fraction of worms growing to adults to below 20%.

20

44. A method as claimed in any one of claims 28 to 43 wherein the step of screening for release from the dauer larval state comprises screening for adult *C. elegans*.

25

45. A method as claimed in any one of claims 28 to 43 wherein said *C. elegans* larvae further comprise a reporter transgene comprising a promoter which is capable of directing strong gene expression in adult *C. elegans* and no or weak expression in dauer larvae or vice versa operably linked to a reporter gene and the step of screening for rescue of the *daf-2* mutation

30

- 82 -

comprises screening for expression of the said reporter gene.

46. A method as claimed in any one of claims 28  
5 to 43 wherein the step of screening for release from the dauer larval state comprises screening for changes in fat storage.

47. A method for the identification of a  
10 compound which is capable of modulating insulin signalling pathways, which method comprises:  
a) providing a sample of nematode worms (preferably eggs, L1 or L2 worms, and most preferably L1 worms);  
15 b) keeping said sample under conditions such, without the presence of any compound(s) to be tested, at least 50%, and preferably at least 60 %, and more preferably at least 70 %, even more preferably at least 80 %, such as 85-100% of the nematodes  
20 present in said sample would enter the dauer state (at least during the time used for the assay);  
c) exposing the sample to the compound(s) to be tested;  
d) measuring either the number of worms that enter the  
25 dauer state, and/or measuring the number of worms that grow into adults.

48. Method according to claim 47, in which the conditions used in step b) are such that, in the  
30 presence of a reference compound at a suitable concentration, the amount of worms that enter the dauer state is at least 10% less, preferably at least

- 83 -

20% less, more preferably at least 30% less, than the amount of worms that would enter the dauer state without the presence of any such reference compound (at least during the time used for the assay).

5

49. Method according to claim 46 and/or 47, in which the conditions used in step b) are such that, in the presence of a reference compound at a suitable concentration, the amount of worms that enter the dauer state is less than 50%, preferably less than 40%, even more preferably less than 30% (at least during the time used for the assay).

50. Method according to any of claims 47-49, in which the nematode worms that form the sample belong to a nematode strain that has an Insulin Sensitivity Value ("ISV") that is greater than the ISV for the reference nematode strain CB1370, in particular more than 1% greater, preferably more than 5% greater, more preferably more than 10% greater, even more preferably more than 20% greater.

51. Method according to any of claims 47-50, in which the nematode worms that form the sample belong to a nematode strain which has an ISV that is >30 %, preferably >40%, even more preferably >50%.

52. Method according to any of claims 47-50, in which the nematodes used in the sample are daf-2(m41) mutants.

53. Use of at least one nematode worm, which has

- 84 -

an increased sensitivity of the insulin signalling pathway, in an assay for the identification of a compound which is capable of modulating insulin signalling pathways.

5

54. Use according to claim 53, in which the nematode worm belongs to a strain that has an Insulin Sensitivity Value ("ISV") that is greater than the ISV for the reference nematode strain CB1370, in particular more than 1% greater, preferably more than 5% greater, more preferably more than 10% greater, even more preferably more than 20% greater.

55. Use according to claim 53 and/or 54, in which the nematode worm belongs to a strain that has an Insulin Sensitivity Value ("ISV") that is >30 %, preferably >40%, even more preferably >50%

56. Use according to any of claims 53-55, in which the nematode worm used is a daf-2(m41) mutant.

57. Use according to any of claims 53-56, in an assay that is carried out in a multi-well plate format.

25

58. Use according to any of claims 53-57, in an assay that is carried out in an automated fashion.

59. Use according to any of claims 53-58, in an assay based on the dauer phenotype as a biological read out, such as on the entry into, the bypass of and/or the rescue from the dauer state, and/or on any

30

- 85 -

other property which results from and/or is associated with the so-called dauer decision.

60. Use according to claim 59, in an assay based  
5 on entry into the dauer state and/or bypass of the  
dauer state as a biological read out.

61. Use according to claim 59, in an assay based  
on rescue from the dauer state as a biological read  
10 out.

62. Use according to any of claims 53-61, for the  
identification of a small molecule and/or a small,  
peptide.  
15

Figure 1: The insulin receptor pathway

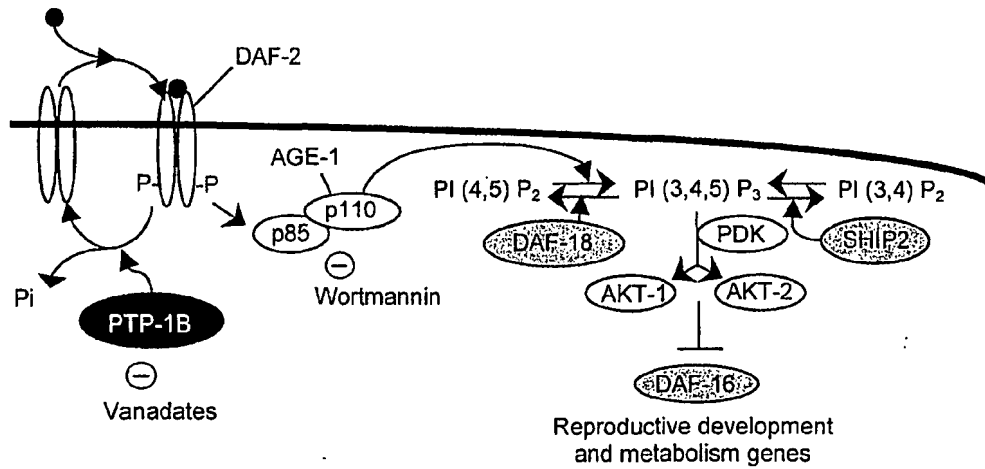


Figure 2. The reference allele of *daf-2* is *e1370*

| Locus: <i>daf-2</i>                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <div> <a href="#">Biblio</a> <a href="#">Attach...</a> <a href="#">Quit</a> </div> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Name                                                                               | Gene class <i>daf</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Type                                                                               | Gene Reference_Allele <b><i>e1370</i></b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Phenotype                                                                          | <i>e1370ts</i> : constitutive dauer formation at 25x; reversible by shift to 15x. ES3 (L3). NR19.<br>See also <i>e1032</i> , <i>e1286</i> , <i>e1365</i> , <i>e1368</i> , <i>e1370</i> , <i>e1391</i><br>[C.elegansIII] <i>e1370ts</i> : constitutive dauer formation at 25C; reversible by shift to 15C. Increased lifespan at 20C; increased thermotolerance, UV resistance, Non-SrP. Synthetic lethal with <i>daf-12</i> , ES3 (L3), OA40; <i>e1032</i> , <i>e1286</i> , <i>e1365</i> , <i>sa230</i> (100%Daf-c at all temperatures), <i>sa223</i> (sterile), <i>m65</i> (nonconditional), etc. Most alleles (not <i>e1370</i> ) hypersensitive to dauer pheromone. [Larsen et al. 1995; Malone and Thomas 1994; CF; JC] |
| Molecular_information                                                              | Sequence <b>EMBL:AF012437.1</b><br><b>EMBL:AF012437.2</b><br><b>Y5505A.391.b</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Map III                                                                            | Position -9,88234 Error 0,059406                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Positive                                                                           | Inside_rearr <b>ndf11</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|                                                                                    | Positive_clone <b>C44811</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|                                                                                    | <b>H05C05</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Negative                                                                           | Outside_rearr <b>tdf9</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Mapping_data                                                                       | Well_ordered                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|                                                                                    | 2_point ----> 4                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|                                                                                    | Multi_point ----> 18                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|                                                                                    | Pos_neg_data ----> 12                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Allele                                                                             | ----> 8                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Strain                                                                             | ----> 13                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Reference                                                                          | ----> 182                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

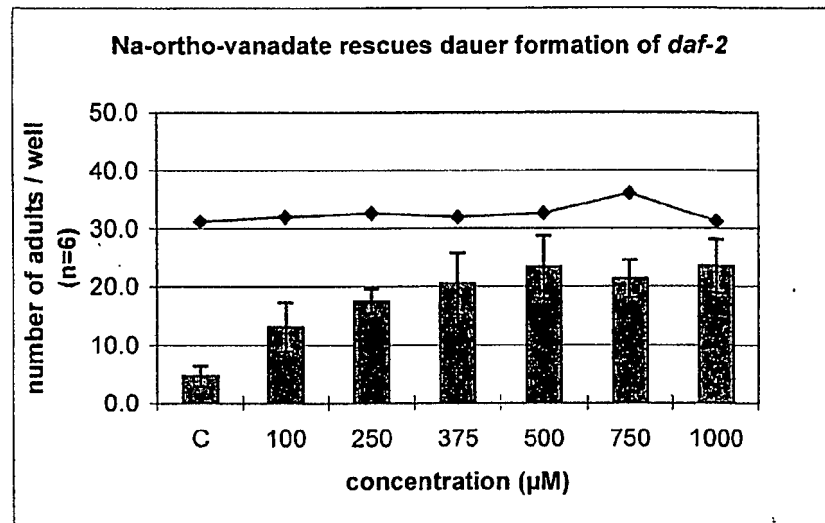
Figure 3: Na-ortho-vanadate rescues insulin resistance caused by *daf-2(m41)*

Figure 4: Wortmannin further enhances insulin resistance

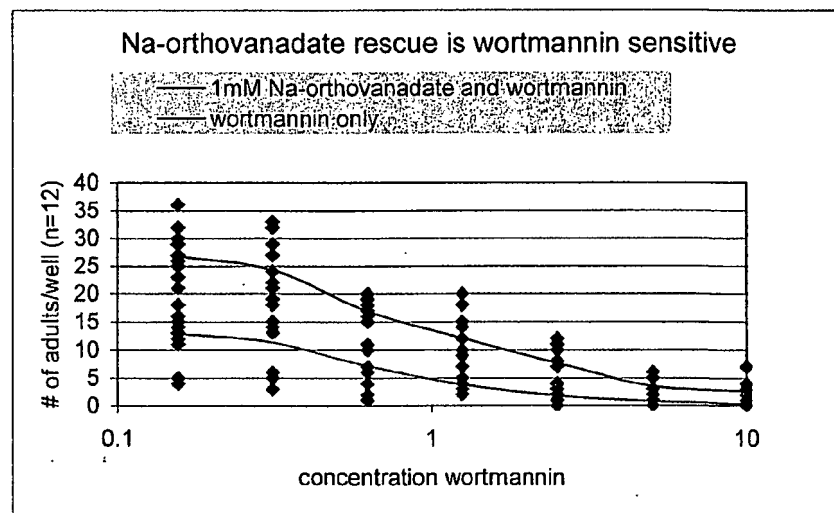


Figure 5: Scatter plots of mean and variance of controls: a (left): screening, b (right): DRC

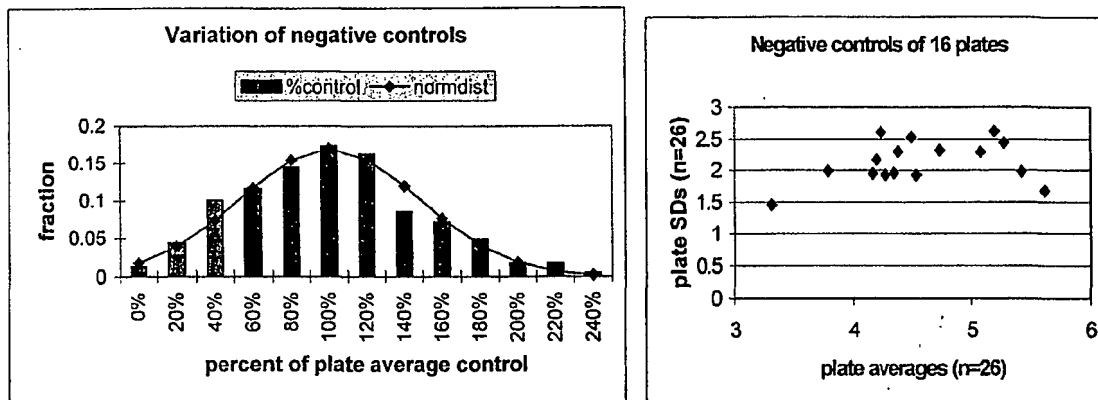


Figure 6: distribution of controls and a maximum likelihood fit of a negative binomial distribution

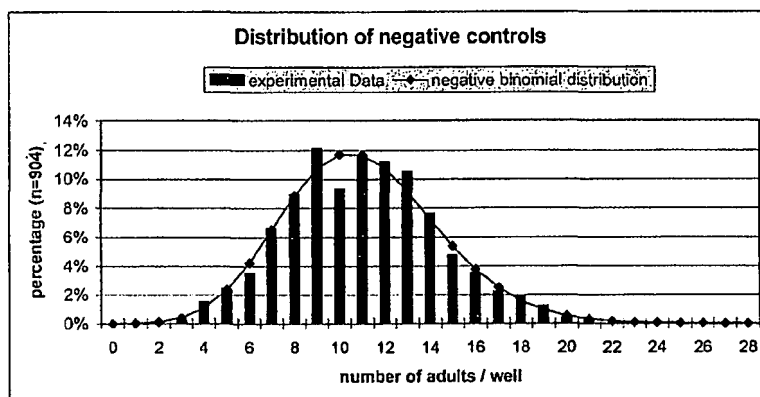


Figure 7: distribution of controls in percent of the average of the plate.

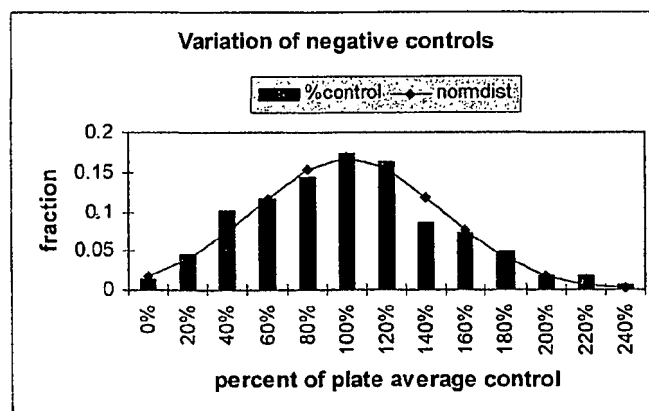


Figure 8

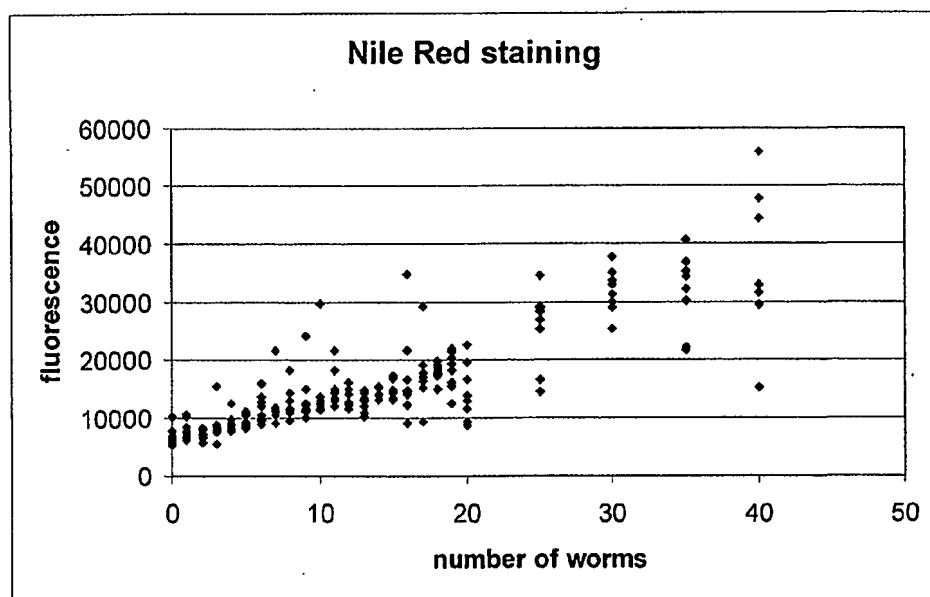


Figure 9

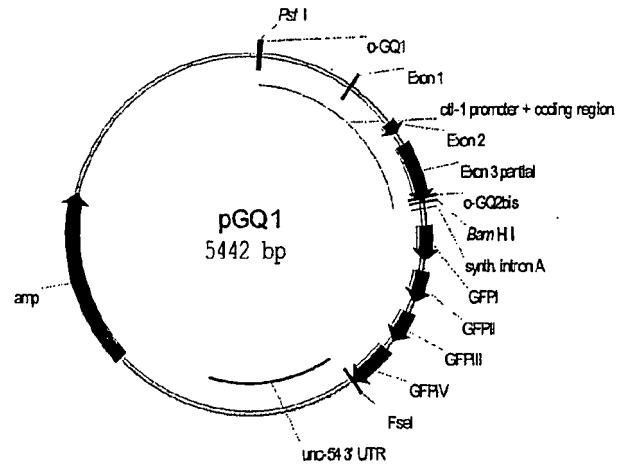


Figure 10

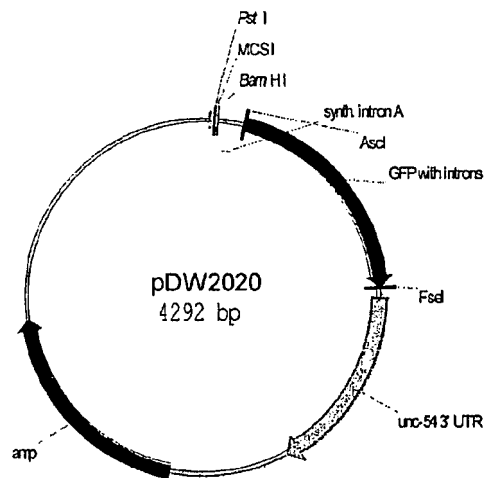


Fig. 11

pDW2020 sequence:

```

 MCS I
 =====
 PstI BamHI
                                     ~~~~~      ~~~
1  ATGACCATGA TTACGCCAAG CTTGCATGCC TGCAGGTCGA CTCTAGAGGA
   TACTGGTACT AATGCGGTTC GAACGTACGG ACGTCCAGCT GAGATCTCCT

MCS I                                synth. intron A
====                                =====
BamHI
~~~

51 TCCCCGGGAT TGGCCAAAGG ACCCAAAGGT ATGTTTCGAA TGATACTAAC
 AGGGGGCCCTA ACCGGTTTCC TGGGTTTCCA TACAAAGCTT ACTATGATTG

synth. intron A
=====

101 ATAACATAGA ACATTTTCAG GAGGACCCTT GGCTAGCGTC GACGGTACCA
 TATTGTATCT TGTAAAAGTC CTCCTGGGAA CCGATCGCAG CTGCCATGGT

AscI GFP with introns
=====

151 TGGGGCGCGC CATGAGTAAA GGAGAAGAAC TTTTCACTGG AGTTGTCCCA
 ACCCGCGCGG GTACTCATT CTCTTCTTG AAAAGTGACC TCAACAGGGT

GFP with introns
=====

201 ATTCTTGTG AATTAGATGG TGATGTTAAT GGGCACAAAT TTTCTGTCAG
 TAAGAACAAC TTAATCTACC ACTACAATTA CCGTGTTTA AAAGACAGTC

GFP with introns
=====

251 TGGAGAGGGT GAAGGTGATG CAACATACGG AAAACTTACC CTAAATTTA
 ACCTCTCCCA CTTCCACTAC GTTGATGCC TTTGAATGG GAATTTAAAT

GFP with introns
=====

301 TTTGCACTAC TGGAAACTA CCTGTTCCAT GGGTAAGTTT AAACATATAT
 AAACGTGATG ACCTTTTGAT GGACAAGGTA CCCATTCAA TTTGTATATA

GFP with introns
=====

351 ATACTAACTA ACCCTGATTA TTTAAATTT CAGCCAACAC TTGTCACCTAC
 TATGATTGAT TGGGACTAAT AAATTTAAAA GTCGGTGTG AACAGTGATG

GFP with introns
=====

401 TTTCTGTTAT GGTGTTCAAT GCTTCTCGAG ATACCCAGAT CATATGAAAC
 AAAGACAATA CCACAAGTTA CGAAGAGCTC TATGGGTCTA GTATACTTTG

GFP with introns
=====

```

fig. 11 continued

451 GGCATGACTT TTTCAAGAGT GCCATGCCCG AAGGTTATGT ACAGGAAAGA  
CCGTACTGAA AAAGTTCTCA CGGTACGGGC TTCCAATACA TGTCCCTTTCT

GFP with introns  
=====

501 ACTATATTTT TCAAAGATGA CGGGAACACTAC AAGACACGTA AGTTTAAACA  
TGATATAAAA AGTTTCTACT GCCCTTGATG TTCTGTGCAT TCAAATTTGT

GFP with introns  
=====

551 GTTCGGTACT AACTAACCAT ACATATTTAA ATTTTCAGGT GCTGAAGTCA  
CAAGCCATGA TTGATTGGTA TGTATAAATT TAAAAGTCCA CGACTTCAGT

GFP with introns  
=====

601 AGTTTGAAGG TGATACCCTT GTTAATAGAA TCGAGTTAA AGGTATTGAT  
TCAAACCTCC ACTATGGGAA CAATTATCTT AGCTCAATTT TCCATAACTA

GFP with introns  
=====

651 TTTAAGAAG ATGGAAACAT TCTTGGACAC AAATTGGAAT ACAACTATAA  
AAATTTCTTC TACCTTTGTA AGAACCTGTG TTAAACCTTA TGTTGATATT

GFP with introns  
=====

701 CTCACACAAT GTATACATCA TGGCAGACAA ACAAAGAAT GGAATCAAAG  
GAGTGTGTTA CATATGTAGT ACCGTCTGTT TGTTTCTTA CCTTAGTTTC

GFP with introns  
=====

751 TTGTAAGTTT AACTTGGAC TTACTAACTA ACGGATTATA TTTAAATTTT  
AACATTCAAA TTTGAACCTG AATGATTGAT TGCCTAATAT AAATTTAAAA

GFP with introns  
=====

801 CAGAACTTCA AAATTAGACA CAACATTGAA GATGGAAGCG TTCAACTAGC  
GTCTTGAAGT TTTAATCTGT GTTGTAACCT CTACCTTCGC AAGTTGATCG

GFP with introns  
=====

851 AGACCATTAT CAACAAAATA CTCCAATTGG CGATGGCCCT GTCCTTTTAC  
TCTGGTAATA GTTGTTTTAT GAGGTTAACC GCTACCGGGA CAGGAAAATG

GFP with introns  
=====

901 CAGACAACCA TTACCTGTCC ACACAATCTG CCCTTTGAA AGATCCCAAC  
GTCTGTTGGT AATGGACAGG TGTGTTAGAC GGGAAAGCTT TCTAGGGTTG

GFP with introns  
=====

951 GAAAAGAGAG ACCACATGGT CCTTCTTGAG TTTGTAACAG CTGCTGGGAT  
CTTTTCTCTC TGGTGTACCA GGAAGAACTC AAACATTGTC GACGACCCTA

GFP with introns

FseI

Fig. 11 continued

```
=====
1001 TACACATGGC ATGGATGAAC TATACAAATA GGGCCGGCCG AGCTCCGCAT
 ATGTGTACCG TACCTACTTG ATATGTTTAT CCCGGCCGGC TCGAGGCGTA
 unc-54 3' UTR
=====
1051 CGGCCGCTGT CATCAGATCG CCATCTCGCG CCCGTGCCTC TGACTTCTAA
 GCCGGCGACA GTAGTCTAGC GGTAGAGCGC GGGCACGGAG ACTGAAGATT
 unc-54 3' UTR
=====
1101 GTCCAATTAC TCTTCAACAT CCCTACATGC TCTTCTCCC TGTGCTCCCA
 CAGGTTAATG AGAAGTTGTA GGGATGTACG AGAAAGAGGG ACACGAGGGT
 unc-54 3' UTR
=====
1151 CCCCTATTT TTGTTATTAT CAAAAAACT TCTTCTAAT TTCTTTGTTT
 GGGGGATAAA AACAATAATA GTTTTTTGA AGAAGAATTA AAGAAACAAA
 unc-54 3' UTR
=====
1201 TTTAGCTTCT TTTAAGTCAC CTCTAACAAT GAAATTGTGT AGATTCAAAA
 AAATCGAAGA AAATTCAGTG GAGATTGTTA CTTTAACACA TCTAAGTTTT
 unc-54 3' UTR
=====
1251 ATAGAATTAA TTCGTAATAA AAAGTCGAAA AAAATTGTGC TCCCTCCCCC
 TATCTTAATT AAGCATTATT TTTCAGCTTT TTTTAACACG AGGGAGGGGG
 unc-54 3' UTR
=====
1301 CATTAATAAT AATTCTATCC CAAAATCTAC ACAATGTTCT GTGTACACTT
 GTAATTATTA TTAAGATAGG GTTTTAGATG TGTTACAAGA CACATGTGAA
 unc-54 3' UTR
=====
1351 CTTATGTTTT TTTTACTTCT GATAAATTT TTTTGAAACA TCATAGAAAA
 GAATACAAAA AAAATGAAGA CTATTTAAAA AAAACTTTGT AGTATCTTTT
 unc-54 3' UTR
=====
1401 AACCGCACAC AAAATACCTT ATCATATGTT ACGTTTCAGT TTATGACCGC
 TTGGCGTGTG TTTTATGGAA TAGTATACAA TGCAAAGTCA AATACTGGCG
 unc-54 3' UTR
=====
1451 AATTTTTATT TCTTCGCACG TCTGGGCCTC TCATGACGTC AAATCATGCT
 TTAAAAATAA AGAAGCGTGC AGACCCGGAG AGTACTGCAG TTTAGTACGA
 unc-54 3' UTR
=====
1501 CATCGTGAAA AAGTTTTGGA GTATTTTGG AATTTTCAA TCAAGTGAAA
 GTAGCACTTT TTCAAACCT CATAAAACC TAAAAAGTT AGTTCACTTT
```

Fig. 11 continued

unc-54 3' UTR  
=====

1551 GTTTATGAAA TTAATTTTCC TGCTTTTGCT TTTTGGGGGT TTCCCTATT  
CAAATACTTT AATTAAAAGG ACGAAAACGA AAAACCCCCA AAGGGGATAA

unc-54 3' UTR  
=====

1601 GTTTGTCAAG AGTTTCGAGG ACGGCGTTTT TCTTGCTAAA ATCACAAGTA  
CAAACAGTTC TCAAAGCTCC TGCCGCAAAA AGAACGATTT TAGTGTTTCAT

unc-54 3' UTR  
=====

1651 TTGATGAGCA CGATGCAAGA AAGATCGGAA GAAGGTTTGG GTTTGAGGCT  
AACTACTCGT GCTACGTTCT TTCTAGCCTT CTCCCAAACC CAAACTCCGA

unc-54 3' UTR  
=====

1701 CAGTGGAAGG TGAGTAGAAG TTGATAATTT GAAAGTGGAG TAGTGTCTAT  
GTCACCTTCC ACTCATCTTC AACTATTAAA CTTTCACCTC ATCACAGATA

unc-54 3' UTR  
=====

1751 GGGGTTTTTG CCTTAAATGA CAGAATACAT TCCCAATATA CCAAACATAA  
CCCCAAAAC GGAATTTACT GTCTTATGTA AGGGTTATAT GGTTCGTTAT

unc-54 3' UTR  
=====

1801 CTGTTTCCTA CTASTCGGCC GTACGGGCCC TTTCGTCTCG CGCGTTTCGG  
GACAAAGGAT GATCAGCCGG CATGCCCGGG AAAGCAGAGC GCGCAAAGCC

1851 TGATGACGGT GAAAACCTCT GACACATGCA GTCCTCCGGAG ACGGTCACAG  
ACTACTGCCA CTTTTGGAGA CTGTGTACGT CGAGGGCCTC TGCCAGTGTC

1901 CTTGTCTGTA AGCGGATGCC GGGAGCAGAC AAGCCCGTCA GGGCGCGTCA  
GAACAGACAT TCGCCTACGG CCCTCGTCTG TTCGGGCAGT CCCGCGCAGT

1951 GCGGGTGTG GCGGGTGTG GGGCTGGCTT AACTATGCGG CATCAGAGCA  
CGCCACAAC CGCCACAGC CCCGACCGAA TTGATACGCC GTAGTCTCGT

2001 GATTGTACTG AGAGTGCACC ATATGCGGTG TGAAATACCG CACAGATGCG  
CTAACATGAC TCTCAGTGG TATACGCCAC ACTTTATGGC GTGTCTACGC

2051 TAAGGAGAAA ATACCGCATC AGGCGGCCTT AAGGGCCTCG TGATACGCCT  
ATTCCTCTTT TATGGCGTAG TCCGCCGGA TTCCCGGAGC ACTATGCGGA

2101 ATTTTTATAG GTTAATGTCA TGATAATAAT GGTTCCTTAG ACGTCAGGTG  
TAAAAATATC CAATTACAGT ACTATTATTA CCAAAGAATC TGCAGTCCAC

2151 GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT ATTTTCTAA  
CGTGAAAAGC CCCTTTACAC GCGCCTTGGG GATAAACAAA TAAAAAGATT

2201 ATACATTCAA ATATGTATCC GTCATGAGA CAATAACCCT GATAAATGCT  
TATGTAAGTT TATACATAGG CGAGTACTCT GTTATTGGGA CTATTTACGA

Fig. 11 continued

```

 amp
=====
2251 TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG
 AGTTATTATA ACTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC

 amp
=====
2301 CCTTATTCC CTTTTTGCG GCATTTGCGC TTCCTGTTTT TGCTCACCCA
 GGGAAATAAGG GAAAAACGC CGTAAAACGG AAGGACAAAA ACGAGTGGGT

 amp
=====
2351 GAAACGCTGG TGAAAGTAAA AGATGCTGAA GATCAGTTGG GTGCACGAGT
 CTTTGCAGACC ACTTTCATTT TCTACGACTT CTAGTCAACC CACGTGCTCA

 amp
=====
2401 GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC
 CCCAATGTAG CTTGACCTAG AGTTGTCGCC ATTCTAGGAA CTCTCAAAG

 amp
=====
2451 GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTAAAGT TCTGCTATGT
 CGGGGCTTCT TGCAAAAGGT TACTACTCGT GAAAATTTC AAGACGATACA

 amp
=====
2501 GGCGCGGTAT TATCCCGTAT TGACGCCGGG CAAGAGCAAC TCGGTCGCGG
 CCGCGCCATA ATAGGGCATA ACTGCGGCC GTTCTCGTTG AGCCAGCGGC

 amp
=====
2551 CATACTAT TCTCAGAATG ACTTGTTGA GTAATCACC GTACAGAAA
 GTATGTGATA AGAGTCTTAC TGAACCAACT CATGAGTGGT CAGTGTCTT

 amp
=====
2601 AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA
 TCGTAGAATG CCTACCGTAC TGTCATTCTC TTAATACGTC ACGACGGTAT

 amp
=====
2651 ACCATGAGTG ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG
 TGGTACTCAC TATTGTGACG CCGGTGAAT GAAGACTGTT GCTAGCCTCC

 amp
=====
2701 ACCGAAGGAG CTAACCGCTT TTTTGCACAA CATGGGGGAT CATGTAACCT
 TGGCTTCCTC GATTGGCGAA AAAACGTGTT GTACCCCTA GTACATTGAG

 amp
=====
2751 GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG
 CGGAACTAGC AACCCTTGCG CTCGACTTAC TTCGGTATGG TTTGCTGCTC
```

Fig. 11. continued

amp  
=====

2801 CGTGACACCA CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACCTATT  
GCACTGTGGT GCTACGGACA TCGTTACCGT TGTGCAACG CGTTTGATAA

amp  
=====

2851 AACTGGCGAA CTACTIONCTC TAGCTTCCCG GCAACAATTA ATAGACTGGA  
TTGACCGCTT GATGAATGAG ATCGAAGGGC CGTTGTTAAT TATCTGACCT

amp  
=====

2901 TGGAGGCGGA TAAAGTTGCA GGACCACTTC TCGCTCGGC CCTTCCGGCT  
ACCTCCGCTT ATTTCAACGT CCTGGTGAAG ACGCGAGCCG GGAAGGCCGA

amp  
=====

2951 GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG  
CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC

amp  
=====

3001 TATCATTGCA GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA  
ATAGTAACGT CGTGACCCCG GTCTACCATT CGGGAGGGCA TAGCATCAAT

amp  
=====

3051 TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC  
AGATGTGCTG CCCCTCAGTC CGTTGATACC TACTTGCTTT ATCTGTCTAG

amp  
=====

3101 GCTGAGATAG GTGCCTCACT GATTAAAGCAT TGCTAACTGT CAGACCAAGT  
CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA GTCTGGTTCA

3151 TTACTCATAT ATACTTTAGA TTGATTAAAA ACTTCATTTT TAATTTAAAA  
AATGAGTATA TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAAATTTT

3201 GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA  
CCTAGATCCA CTCTCTAGGAA AAACATATTAG AGTACTGGTT TTAGGGAATT

3251 CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG  
GCACTCAAAA GCAAGGTGAC TCGCAGTCTG GGGCATCTTT TCTAGTTTCC

3301 ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA  
TAGAAGAACT CTAGGAAAAA AAGACGCGCA TTAGACGACG AACGTTTGTT

3351 AAAAACCACC GCTACCAGCG GTGGTTTGTT TGCCGGATCA AGAGCTACCA  
TTTTTGGTGG CGATGGTCGC CACCAAACAA ACGGCCTAGT TCTCGATGCT

3401 ACTCTTTTTT CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC  
TGAGAAAAAG GCTTCCATTG ACCGAAGTCG TCTCGCGTCT ATGGTTTATG

Fig. 11 continued

3451 TGTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG  
 ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC  
 3501 CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC  
 GTGGCGGATG TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG  
 3551 AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC  
 TCACCGCTAT TCAGCACAGA ATGGCCCAAC CTGAGTTCTG CTATCAATGG  
 3601 GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA  
 CCTATTCCGC GTCGCCAGCC CGACTTGCCC CCAAGCACG TGTGTCGGGT  
 3651 GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCAT  
 CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGTA  
 3701 TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT  
 ACTCTTTTCGC GGTGCGAAGG GCTTCCCTCT TTCCGCCTGT CCATAGGCCA  
 3751 AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA  
 TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG CTCCTCGAA GGTCCCCCTT  
 3801 ACGCCTGGTA TCTTTATAGT CCTGTGCGGT TTCGCCACCT CTGACTTGAG  
 TGCGGACCAT AGAAATATCA GGACAGCCCA AAGCGGTGGA GACTGAACTC  
 3851 CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC  
 CGAGCTAAAA ACACTACGAG CAGTCCCCC GCCTCGGATA CCTTTTTCGC  
 3901 CAGCAACGCG GCCTTTTAC GGTTCCTGGC CTTTGTGCTG CCTTTTGCTC  
 GTCGTTGCCG CGGAAAAATG CCAAGGACCG GAAAACGACC GAAAACGAG  
 3951 ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA CCGTATTACC  
 TGTACAAGAA AGGACGCAAT AGGGGACTAA GACACCTATT GGCATAATGG  
 4001 GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG  
 CGGAAACTCA CTCGACTATG GCGAGCGGCG TCGGCTTGCT GGCTCGCGTC  
 4051 CGAGTCAGTG AGCGAGGAAG CGGAAGAGCG CCCAATACGC AAACCGCCTC  
 GCTCAGTCAC TCGCTCCTTC GCCTTCTCGC GGGTTATGCG TTTGGCGGAG  
 4101 TCCCCGCGCG TTGGCCGATT CATTAATGCA GCTGGCACGA CAGGTTTCCC  
 AGGGGCGCGC AACC GGCTAA GTAATTACGT CGACCGTGCT GTCCAAAGGG  
 4151 GACTGGAAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA GTTAGCTCAC  
 CTGACCTTTC GCCCGTCACT CGCGTTGCGT TAATTACACT CAATCGAGTG  
 4201 TCATTAGGCA CCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT  
 AGTAATCCGT GGGGTCCGAA ATGTGAAATA CGAAGGCCGA GCATACAACA  
 4251 GTGGAATTGT GAGCGGATAA CAATTTACA CAGGAAACAG CT  
 CACCTTAACA CTCGCCTATT GTTAAAGTGT GTCCTTTGTC GA

Fig. 12

## II. Predicted DNA sequence pGQ1

```

 ctl-1 promoter + coding region
 =
 o-GQ1
 =
 PstI
                                ~~~~~
1  ATGACCATGA TTACGCCAAG CTTGCATGCC TGCAGCCAAT GCATTGGAAG
   TACTGGTACT AATGCGGTTC GAACGTACGG ACGTCGGTTA CGTAACCTTC

                                ctl-1 promoter + coding region
=====
                                o-GQ1
=====
51 AGATATTTTG CGCGTCAAAT ATGTTTGTG TCCCGTAAT ATTTTTTTAA
   TCTATAAAAC GCGCAGTTTA TACAAAACAC AGGGGCATTA TAAAAAATT

                                ctl-1 promoter + coding region
=====
101 ATCAAATTC ACATTTTAAC CATAAAAAAC TCTTCAAAA GTGTAATTTT
   TAGTTTAAAG TGTAATAATTG GTATTTTTTG AGAAAGTTT CACATTAAAA

                                ctl-1 promoter + coding region
=====
151 CTACGCAAAA ATGCCGTCG GATGAAAAAT TACTTTTGAA AAACAACTC
   GATGCGTTTT TACGGCAAGC CTACTTTTTA ATGAAACTT TTTGTTGAG

                                ctl-1 promoter + coding region
=====
201 GAAACTACGG TACGCAAAAA AGTACATCGG TGTTCGACA TAAGTGAAAA
   CTTTGATGCC ATGCGTTTTT TCATGTAGCC ACAAACGTGT ATTCACTTT

                                ctl-1 promoter + coding region
=====
251 CAATGTTGTT TTTTGTAAT TAAATCGAT TAATTTTTTT TCCCGGAAAA
   GTTACAACAA AAAACATTA ATTTAGCTA ATTAAAAAAA AGGGCCTTT

                                ctl-1 promoter + coding region
=====
301 CAAAAACGTT TTCAGCGTGG ATTTCTATTG TTTCTGCGT AAAAAAAAT
   GTTTTGGCAA AAGTCGCACC TAAAGATAAC AAAGAACGCA TTTTTTTTA

                                ctl-1 promoter + coding region
=====
351 TATTTACCAA TTTTAAACGA TAATTCCAC GAATTTTCGC CATTAACTC
   ATAAATGGTT AAAATTGCT ATTAAAGGTG CTAAAAGCG GTAATTAGAG

                                ctl-1 promoter + coding region
=====
401 TCGATTTTGT TGATTCTTGA CTCCGAGCAA TCTCTCCGT TTCGCAAA
   AGCTAAAACA ACTAAGAACT GAGGCTCGTT AGAGAGGCCA AAAGCGTTG

```

Fig. 12 continued

```

                                ctl-1 promoter + coding region
                                =====
451  GATTATATTA TTTATTTGTT TTCCTTTTCA GTGCCGATTC TCGGAAATTC
    CTAATATAAT AAATAAACAA AAGGAAAAGT CACGGCTAAG AGCCTTTAAG

                                ctl-1 promoter + coding region
                                =====
                                Exon 1
                                =====
501  AACAGTAAAT CTTCAAAATG CCAATGCTTC CCCACATGGT CAATCTAAGT
    TTGTCATTTA GAAGTTTTAC GGTACGAAG GGGTGTACCA GTTAGATTCA

                                ctl-1 promoter + coding region
                                =====
551  GAGTTTCTTT GTTACAAAAT ACACGTGATG TCAGATTGTC TCATTTTCGGT
    CTCAAAGAAA CAATGTTTTA TGTGCACTAC AGTCTAACAG AGTAAAGCCA

                                ctl-1 promoter + coding region
                                =====
601  TTGATCTACG TAGATCTACA AAAAATGCGG GAATTGAGCC GCAGAGTTCT
    AACTAGATGC ATCTAGATGT TTTTACGCC CTTAACTCGG CGTCTCAAGA

                                ctl-1 promoter + coding region
                                =====
651  CAACTGCTTT CGCATGGTTA AGAACGTGCG GACGTCAAAT TGTTTTGGGC
    GTTGACGAAA GCGTACCAAT TCTTGACGCG CTGCAGTTTA ACAAACCCG

                                ctl-1 promoter + coding region
                                =====
701  AAAAAATCCC GCATTTTTTG TAGATCAAAC CGTAATGGGA CAGTCTGGCA
    TTTTAAAGGG CGTAAAAAAC ATCTAGTTTG GCATTACCCT GTCAGACCGT

                                ctl-1 promoter + coding region
                                =====
                                Exon 2
                                =====
751  CCACGTGACT ATATATTTTT AGCGGTCAAC GACACAAAAC CCGGACCAAT
    GGTGCACTGA TATATAAAAA TCGCCAGTTG CTGTGTTTTG GGCCTGGTTA

                                ctl-1 promoter + coding region
                                =====
                                Exon 2
                                =====
801  GGCTGAGGAT CAGCTGAAAG CTTATAGAGA TAGAAATCAG GTGAGAAAAA
    CCGACTCCTA GTCGACTTTC GAATATCTCT ATCTTTAGTC CACTCTTTTT

                                ctl-1 promoter + coding region
                                =====
851  TCAATTCAG CGATTTTCTT CGCAATTAT ATAAAACTG ATTTTCCAG
    AGTTAAAGTC GCTAAAAGAA GCGTTAAATA TATTTTGAC TAAAAAGGTC

                                ctl-1 promoter + coding region
                                =====
                                Exon 3 partial

```

Fig. 12 continued

901 GAACCCACACC TGCTCACCAC ATCCAATGGA GCTCCGATCT ACTCGAAGAC  
CTTGGGGTGG ACGAGTGGTG TAGGTTACCT CGAGGCTAGA TGAGCTTCTG

ctl-1 promoter + coding region

Exon 3 partial

951 CGCCGTGCTC ACCGCCGGAC GACGTGGTCC AATGCTAATG CAGGACATCG  
GCGGCACGAG TGGCGGCCTG CTGCACCAGG TTACGATTAC GTCCTGTAGC

ctl-1 promoter + coding region

Exon 3 partial

1001 TTTATATGGA CGAGATGGCT CATTTCGATC GTGAACGCAT CCCGGAGCGT  
AAATATACCT GCTCTACCGA GTAAAGCTAG CACTTGCGTA GGGCCTCGCA

ctl-1 promoter + coding region

Exon 3 partial

1051 GTCGTCCATG CCAAAGGTGG TGGTGCTCAT GGATACTTCG AGGTCACCCA  
CAGCAGGTAC GGTTCACC ACCACGAGTA CCTATGAAGC TCCAGTGGGT

ctl-1 promoter + coding region

Exon 3 partial

1101 TGACATCACC AAGTACTGTA AGGCCGATAT GTTCAACAAG GTCGGAAAAC  
ACTGTAGTGG TTCATGACAT TCCGGCTATA CAAGTTGTTC CAGCCTTTTG

ctl-1 promoter + coding region

o-GQ2bis

Exon 3 partial

BamHI

1151 AGACACCACT TCTCGTTCGT TTTTCAACGG TCGCTGGAGA ATCGGCCGGA  
TCTGTGGTGA AGAGCAAGCA AAAAGTTGCC AGCGACCTCT TAGCCGGCCT

ctl-1 promoter + coding region

o-GQ2bis

Exon 3 partial

synth. intron A

BamHI

1201 TCCCCGGGAT TGGCCAAAGG ACCCAAAGGT ATGTTTCGAA TGATACTAAC  
AGGGGCCCTA ACCGGTTTCC TGGGTTTCCA TACAAAGCTT ACTATGATTG

Fig. 12 continued

synth. intron A

1251 ATAACATAGA ACATTTTCAG GAGGACCCTT GGCTAGCGTC GACGGTACCA  
TATTGTATCT TGTAAAAGTC CTCCTGGGAA CCGATCGCAG CTGCCATGGT

GFPI

1301 TGGGGCGCGC CATGAGTAAA GGAGAAGAAC TTTTCACTGG AGTTGTCCCA  
ACCCGCGCGG GTACTCATTT CCTCTTCTTG AAAAGTGACC TCAACAGGGT

GFPI

1351 ATTCTTGTTG AATTAGATGG TGATGTTAAT GGGCACAAAT TTTCTGTCAG  
TAAGAACAAC TTAATCTACC ACTACAATTA CCCGTGTTTA AAAGACAGTC

GFPI

1401 TGGAGAGGGT GAAGGTGATG CAACATACGG AAAACTTACC CTAAATTTA  
ACCTCTCCCA CTTCCACTAC GTTGATGCC TTTGAATGG GAATTTAAAT

GFPI

1451 TTTGCACTAC TGGAAACTA CCTGTTCAT GGGTAAGTTT AAACATATAT  
AAACGTGATG ACCTTTTGAT GGACAAGGTA CCCATTCAA TTTGTATATA

GFPII

1501 ATACTAACTA ACCCTGATTA TTAAATTTT CAGCCAACAC TTGTCCTAC  
TATGATTGAT TGGGACTAAT AAATTTAAAA GTCGGTTGTG AACAGTGATG

GFPII

1551 TTTCTGTTAT GGTGTTCAAT GCTTCTCGAG ATACCCAGAT CATATGAAAC  
AAAGACAATA CCACAAGTTA CGAAGAGCTC TATGGGTCTA GTATACTTG

GFPII

1601 GGCATGACTT TTTCAAGAGT GCCATGCCCG AAGGTTATGT ACAGGAAAGA  
CCGTACTGAA AAAGTTCTCA CGGTACGGGC TTCCAATACA TGTCTTTCT

GFPII

1651 ACTATATTTT TCAAAGATGA CGGGAACACTAC AAGACACGTA AGTTTAAACA  
TGATATAAAA AGTTTCTACT GCCCTTGATG TTCTGTGCAT TCAAATTTGT

GFPIII

1701 GTTCGGTACT AACTAACCAT ACATATTTAA ATTTTCAGGT GCTGAAGTCA  
CAAGCCATGA TTGATTGGTA TGTATAAATT TAAAAGTCCA CGACTTCAGT

GFPIII

1751 AGTTTGAAGG TGATACCCTT GTTAATAGAA TCGAGTTAAA AGGTATTGAT  
TCAAACCTTC ACTATGGGAA CAATTATCTT AGCTCAATTT TCCATAACTA

Fig. 12 continued

GFPIII  
=====

1801 TTTAAAGAAG ATGGAAACAT TCTTGGACAC AAATTGGAAT ACAACTATAA  
AAATTTCTTC TACCTTTGTA AGAACCTGTG TTAAACCTTA TGTGTGATATT

GFPIII  
=====

1851 CTCACACAAT GTATACATCA TGGCAGACAA ACAAAGAAT GGAATCAAAG  
GAGTGTGTTA CATATGTAGT ACCGTCTGTT TGTTTCTTA CCTAGTTTC

GFPIII  
==

1901 TTGTAAGTTT AAAGTTGGAC TTACTAATA ACGGATTATA TTTAAATTTT  
AACATTCAA TTTGAACCTG AATGATTGAT TGCCTAATAT AAATTTAAAA

GFPIV  
=====

1951 CAGAACTTCA AAATTAGACA CAACATTGAA GATGGAAGCG TTCACTAGC  
GTCTTGAAGT TTTAATCTGT GTTGTAACTT CTACCTTCGC AAGTTGATCG

GFPIV  
=====

2001 AGACCATTAT CAACAAAATA CTCCAATTGG CGATGGCCCT GTCCTTTTAC  
TCTGGTAATA GTTGTTTTAT GAGGTTAACC GCTACCGGGA CAGGAAAATG

GFPIV  
=====

2051 CAGACAACCA TTACCTGTCC ACACAATCTG CCCTTTCGAA AGATCCCAAC  
GTCTGTTGGT AATGGACAGG TGTGTTAGAC GGGAAAGCTT TCTAGGGTTG

GFPIV  
=====

2101 GAAAAGAGAG ACCACATGGT CCTTCTTGAG TTTGTAAACAG CTGCTGGGAT  
CTTTTCTCTC TGGTGTACCA GGAAGAACTC AAACATTGTC GACGACCCTA

GFPIV FseI  
=====

2151 TACACATGGC ATGGATGAAC TATACAAATA GGGCCGGCCG AGCTCCGCAT  
ATGTGTACCG TACCTACTTG ATATGTTTAT CCCGGCCGGC TCGAGGCGTA

unc-54 3' UTR  
=====

2201 CGGCCGCTGT CATCAGATCG CCATCTCGCG CCCGTGCCTC TGAATTCTAA  
GCCGGCGACA GTAGTCTAGC GGTAGAGCGC GGGCACGGAG ACTGAAGATT

unc-54 3' UTR  
=====

2251 GTCCAATTAC TCTTCAACAT CCCTACATGC TCTTTCTCCC TGTGCTCCCA  
CAGGTTAATG AGAAGTTGTA GGGATGTACG AGAAAGAGGG ACACGAGGGT

unc-54 3' UTR  
=====

2301 CCCCTATTT TTGTTATTAT CAAAAAACT TCTTCTTAAT TTCTTTGTTT

Fig. 12 continued

GGGGGATAAA AACAAATAATA GTTTTTTTGA AGAAGAATTA AAGAAACAAA  
unc-54 3' UTR  
=====

2351 TTTAGCTTCT TTTAAGTCAC CTCTAACAAT GAAATTGTGT AGATTCAAAA  
AAATCGAAGA AAATTCAGTG GAGATTGTTA CTTTAACACA TCTAAGTTTT  
unc-54 3' UTR  
=====

2401 ATAGAATTAA TTCGTAATAA AAAGTCGAAA AAAATTGTGC TCCCTCCCCC  
TATCTTAATT AAGCATTATT TTTCAGCTTT TTTTAACACG AGGGAGGGGG  
unc-54 3' UTR  
=====

2451 CATTAATAAT AATTCTATCC CAAAATCTAC ACAATGTTCT GTGTACACTT  
GTAATTATTA TTAAGATAGG GTTTTAGATG TGTTACAAGA CACATGTGAA  
unc-54 3' UTR  
=====

2501 CTTATGTTTT TTTTACTTCT GATAAATTTT TTTTGAAACA TCATAGAAAA  
GAATACAAAA AAAATGAAGA CTATTTAAAA AAAACTTTGT AGTATCTTTT  
unc-54 3' UTR  
=====

2551 AACCGCACAC AAAATACCTT ATCATATGTT ACGTTTCAGT TTATGACCGC  
TTGGCGTGTG TTTTATGGAA TAGTATACAA TGCAAAGTCA AATACTGGCG  
unc-54 3' UTR  
=====

2601 AATTTTTATT TCTTCGCACG TCTGGGCCTC TCATGACGTC AAATCATGCT  
TTAAAAATAA AGAAGCGTGC AGACCCGGAG AGTACTGCAG TTTAGTACGA  
unc-54 3' UTR  
=====

2651 CATCGTGAAA AAGTTTTGGA GTATTTTGG AATTTTCAA TCAAGTGAAA  
GTAGCACTTT TTCAAAACCT CATAAAACCC TAAAAAGTT AGTTCACTTT  
unc-54 3' UTR  
=====

2701 GTTTATGAAA TTAATTTTCC TGCTTTTGCT TTTTGGGGGT TTCCCCTATT  
CAAATACTTT AATTAAAAGG ACGAAAACGA AAAACCCCCA AAGGGGATAA  
unc-54 3' UTR  
=====

2751 GTTTGTCAAG AGTTTCGAGG ACGGCGTTTT TCTTGCTAAA ATCACAAGTA  
CAAACAGTTC TCAAAGCTCC TGCCGCAAAA AGAACGATT TAGTGTTTCAT  
unc-54 3' UTR  
=====

2801 TTGATGAGCA CGATGCAAGA AAGATCGGAA GAAGGTTTGG GTTTGAGGCT  
AACTACTCGT GCTACGTTCT TTCTAGCCTT CTCCAAACC CAACTCCGA  
unc-54 3' UTR  
=====

Fig. 12 continued

2851 CAGTGGGAAGG TGAGTAGAAG TTGATAATTT GAAAGTGGAG TAGTGTCTAT  
GTCACCTTCC ACTCATCTTC AACTATTAAA CTTTCACCTC ATCACAGATA

unc-54 3' UTR  
=====

2901 GGGGTTTTTG CTTTAAATGA CAGAATACAT TCCCAATATA CCAAACATAA  
CCCCAAAAAC GGAATTTACT GTCTTATGTA AGGGTTATAT GGTTTGTATT

unc-54 3' UTR  
=====

2951 CTGTTTCCTA CTAGTCGGCC GTACGGGCCC TTTCGTCTCG CGCGTTTCGG  
GACAAAGGAT GATCAGCCGG CATGCCCGGG AAAGCAGAGC GCGCAAAGCC

3001 TGATGACGGT GAAAACCTCT GACACATGCA GCTCCCGGAG ACGGTCACAG  
ACTACTGCCA CTTTTGGAGA CTGTGTACGT CGAGGGCCTC TGCCAGTGTC

3051 CTTGTCTGTA AGCGGATGCC GGGAGCAGAC AAGCCCGTCA GGGCGCGTCA  
GAACAGACAT TCGCCTACGG CCTCGTCTG TTCGGGCAGT CCCGCGCAGT

3101 GCGGGTGTTG GCGGGTGTCG GGGCTGGCTT AACTATGCGG CATCAGAGCA  
CGCCCACAAC CGCCCACAGC CCCGACCGAA TTGATACGCC GTAGTCTCGT

3151 GATTGTACTG AGAGTGCACC ATATGCGGTG TGAAATACCG CACAGATGCG  
CTAACATGAC TCTCAGTGG TATACGCCAC ACTTTATGGC GTGTCTACGC

3201 TAAGGAGAAA ATACCGCATC AGGCGGCCTT AAGGGCCTCG TGATACGCCT  
ATTCTCTTT TATGGCGTAG TCCGCCGAA TCCCGGAGC ACTATGCGGA

3251 ATTTTTATAG GTTAATGTCA TGATAATAAT GGTTTCTTAG ACGTCAGGTG  
TAAAAATATC CAATTACAGT ACTATTATTA CCAAAGAATC TGCAGTCCAC

3301 GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT ATTTTTCTAA  
CGTGAAAAGC CCCTTTACAC GCGCCTTGGG GATAAACAAA TAAAAAGATT

3351 ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT  
TATGTAAGTT TATACATAGG CGAGTACTCT GTTATTGGGA CTATTTACGA

amp  
=====

3401 TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCTG  
AGTTATTATA ACTTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC

amp  
=====

3451 CCCTTATTC CTTTTTTCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA  
GGGAATAAGG GAAAAACGC CGTAAAACGG AAGGACAAA ACGAGTGGGT

amp  
=====

3501 GAAACGCTGG TGAAAGTAAA AGATGCTGAA GATCAGTTGG GTGCACGAGT  
CTTGCGACC ACTTTCATTT TCTACGACTT CTAGTCAACC CACGTGCTCA

amp  
=====

Fig. 12 continued

3551 GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC  
CCCAATGTAG CTTGACCTAG AGTTGTCGCC ATTCTAGGAA CTCTCAAAAG  
amp  
=====

3601 GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTAAAGT TCTGCTATGT  
CGGGGCTTCT TGCAAAAGGT TACTACTCGT GAAAATTTCA AGACGATACA  
amp  
=====

3651 GGCGCGGTAT TATCCCGTAT TGACGCCGGG CAAGAGCAAC TCGGTCGCCG  
CCGCGCCATA ATAGGGCATA ACTGCGGCCG GTTCTCGTTG AGCCAGCGGC  
amp  
=====

3701 CATACTAT TCTCAGAATG ACTTGGTTGA GTACTACCA GTCACAGAAA  
GTATGTGATA AGAGTCTTAC TGAACCACT CATGAGTGGT CAGTGTCTTT  
amp  
=====

3751 AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA  
TCGTAGAATG CCTACCGTAC TGCATTCTC TTAATACGTC ACGACGGTAT  
amp  
=====

3801 ACCATGAGTG ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG  
TGGTACTCAC TATTGTGACG CCGGTTGAAT GAAGACTGTT GCTAGCCTCC  
amp  
=====

3851 ACCGAAGGAG CTAACCGCTT TTTTGACAA CATGGGGGAT CATGTAATC  
TGGCTTCCTC GATTGGCGAA AAAACGTGTT GTACCCCTA GTACATTGAG  
amp  
=====

3901 GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG  
CGGAACCTAGC AACCCCTGGC CTCGACTTAC TTCGGTATGG TTTGCTGCTC  
amp  
=====

3951 CGTGACACCA CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAATATT  
GCACTGTGGT GCTACGGACA TCGTTACCGT TGTGCAACG CGTTTGATAA  
amp  
=====

4001 AACTGGCGAA CTAATTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA  
TTGACCGCTT GATGAATGAG ATCGAAGGGC CGTTGTTAAT TATCTGACCT  
amp  
=====

4051 TGGAGGCGGA TAAAGTTGCA GGACCACTC TCGCTCGGC CCTTCCGGCT  
ACCTCCGCCT ATTTCAACGT CCTGGTGAAG ACGCGAGCCG GGAAGGCCGA  
amp

Fig. 12 continued

```
=====
4101 GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG
    CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC

    amp
    =====
4151 TATCATTGCA GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA
    ATAGTAACGT CGTGACCCCG GTCTACCATT CGGGAGGGCA TAGCATCAAT

    amp
    =====
4201 TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC
    AGATGTGCTG CCCCTCAGTC CGTTGATACC TACTTGCTTT ATCTGTCTAG

    amp
    =====
4251 GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT
    CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA GTCTGGTTCA

4301 TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA
    AATGAGTATA TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAATTTT

4351 GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA
    CCTAGATCCA CTTCTAGGAA AACTATTAG AGTACTGGTT TTAGGGAATT

4401 CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG
    GCACTCAAAA GCAAGGTGAC TCGCAGTCTG GGGCATCTTT TCTAGTTTCC

4451 ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA
    TAGAAGAACT CTAGGAAAAA AAGACGCGCA TTAGACGACG AACGTTTGTT

4501 AAAAACCACC GCTACCAGCG GTGGTTTGTT TGCCGGATCA AGAGCTACCA
    TTTTTTGGTG CGATGGTCGC CACCAAACAA ACGGCCTAGT TCTCGATGGT

4551 ACTCTTTTTT CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC
    TGAGAAAAAG GCTTCCATTG ACCGAAGTCG TCTCGCGTCT ATGGTTTATG

4601 TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG
    ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC

4651 CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC
    GTGGCGGATG TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG

4701 AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC
    TCACCGCTAT TCAGCACAGA ATGGCCCAAC CTGAGTTCTG CTATCAATGG

4751 GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA
    CCTATTCCGC GTCGCCAGCC CGACTTGCCC CCAAGCACG TGTGTCGGGT

4801 GCTTGAGCGC AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCAT
    CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGTA

4851 TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT
    ACTCTTTCGC GGTGCGAAGG GCTTCCCTCT TTCCGCCTGT CCATAGGCCA
```

Fig. 12 continued

4901 AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA  
TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG CTCCCTCGAA GTCCCCCTT

4951 ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG  
TGCGGACCAT AGAAATATCA GGACAGCCCA AAGCGGTGGA GACTGAACTC

5001 CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC  
GCAGCTAAAA ACACTACGAG CAGTCCCCC GCCTCGGATA CCTTTTTCG

5051 CAGCAACGCG GCCTTTTAC GGTTCCTGGC CTTTGTCTGG CCTTTTGCTC  
GTCGTTGCGC CGGAAAAATG CCAAGGACCG GAAACGACC GGAAACGAG

5101 ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA CCGTATTACC  
TGTACAAGAA AGGACGCAAT AGGGGACTAA GACACCTATT GGCATAATGG

5151 GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG  
CGGAAACTCA CTCGACTATG GCGAGCGGCG TCGGCTTGCT GGCTCGCGTC

5201 CGAGTCAGTG AGCGAGGAAG CGGAAGAGCG CCCAATACGC AAACCGCCTC  
GTCAGTCAC TCGCTCCTC GCCTTCTCGC GGGTTATGCG TTTGGCGGAG

5251 TCCCCGCGCG TTGGCCGATT CATTAATGCA GCTGGCACGA CAGGTTTCCC  
AGGGGCGCGC AACCGGCTAA GTAATTACGT CGACCGTGCT GTCCAAAGGG

5301 GACTGGAAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA GTTAGCTCAC  
CTGACCTTTC GCCCGTCACT CGCGTTGCGT TAATTACACT CAATCGAGTG

5351 TCATTAGGCA CCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT  
AGTAATCCGT GGGGTCCGAA ATGTGAAATA CGAAGGCCGA GCATACAACA

5401 GTGGAATTGT GAGCGGATAA CAATTTACA CAGGAAACAG CT  
CACCTTAACA CTCGCCTATT GTTAAAGTGT GTCCTTTGTC GA

Fig. 13

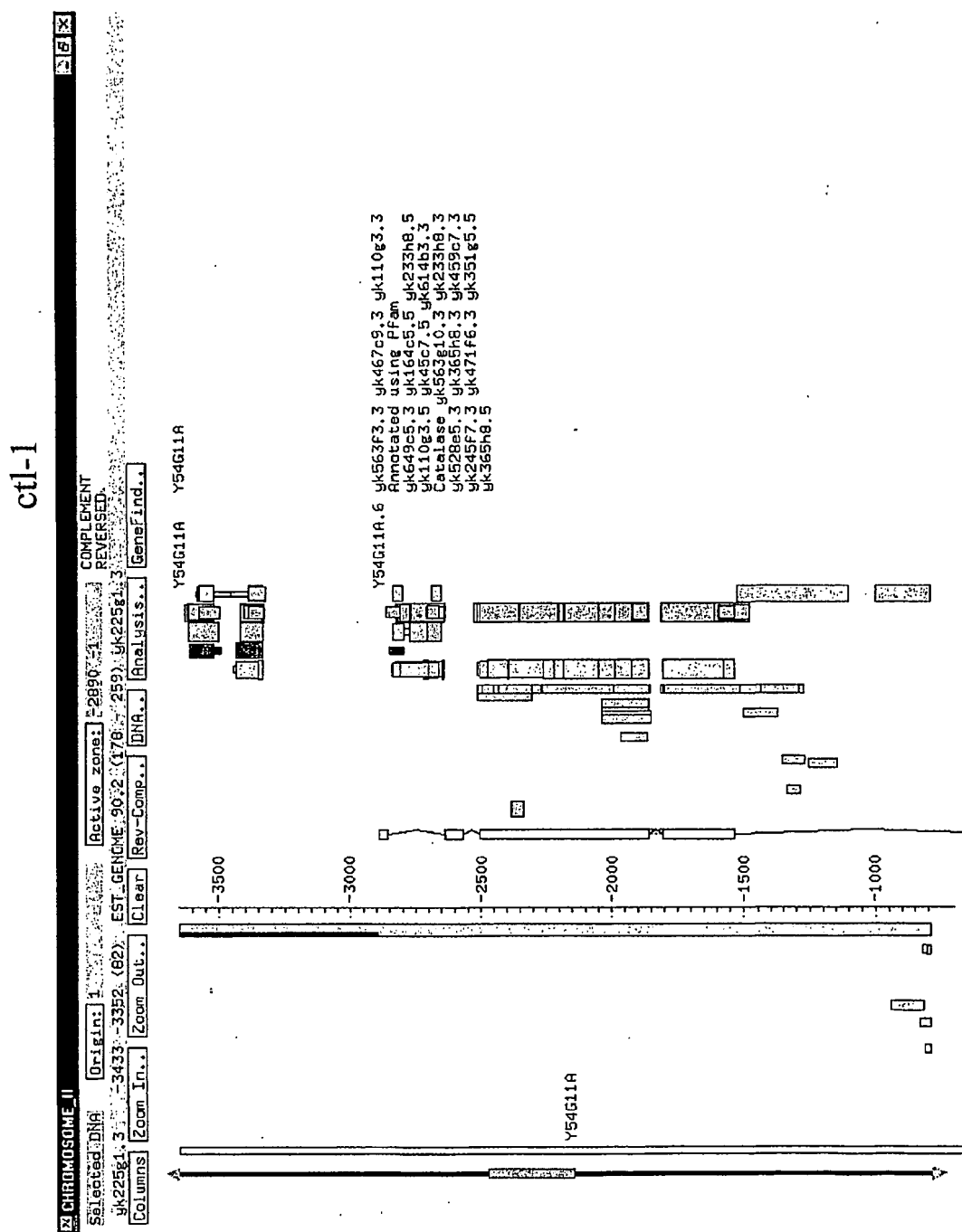


Figure 14

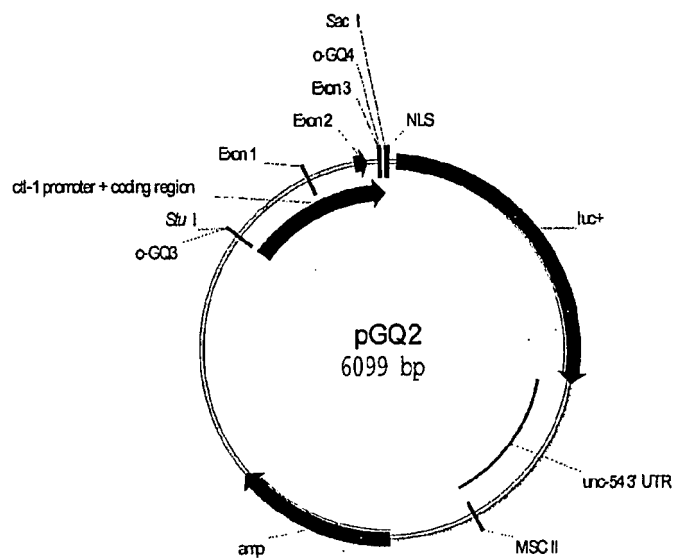


Figure 15

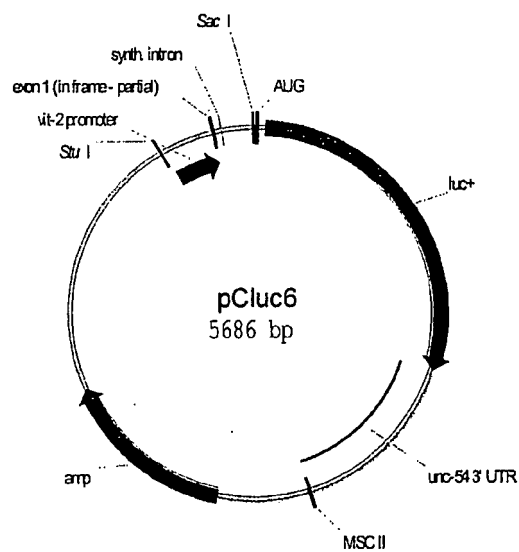


Fig. 16

pCluc6 sequence:

```

      AUG                                     luc+
      ===                                     =====
1   ATGACTGCTC CAAAGAAGAA GCGTAAGGTA CCGGTAGAAA AAATGGAAGA
   TACTGACGAG GTTCTTTCTT CGCATCCAT GGCCATCTTT TTTACCTTCT

                                     luc+
                                     =====
51  CGCCAAAAC ATAAAGAAAG GCCCGGCGCC ATTCTATCCG CTGGAAGATG
   GCGGTTTTTG TATTTCTTTC CGGGCCGCGG TAAGATAGGC GACCTTCTAC

                                     luc+
                                     =====
101 GAACCGCTGG AGAGCAACTG CATAAGGCTA TGAAGAGATA CGCCTTGGTT
   CTTGGCGACC TCTCGTTGAC GTATTCCGAT ACTTCTCTAT GCGGGACCAA

                                     luc+
                                     =====
151 CCTGGAACAA TTGCTTTTAC AGATGCACAT ATCGAGGTGG ACATCACTTA
   GGACCTTGTT AACGAAAATG TCTACGTGTA TAGCTCCACC TGTAGTGAAT

                                     luc+
                                     =====
201 CGCTGAGTAC TTCGAAATGT CCGTTCGGTT GGCAGAAGCT ATGAAACGAT
   GCGACTCATG AAGCTTTACA GGCAAGCCAA CCGTCTTCGA TACTTTGCTA

                                     luc+
                                     =====
251 ATGGGCTGAA TACAAATCAC AGAATCGTCG TATGCAGTGA AAACCTCTCT
   TACCCGACTT ATGTTTAGTG TCTTAGCAGC ATACGTCAC TTTGAGAGAA

                                     luc+
                                     =====
301 CAATTCTTTA TGCCGGTGTT GGGCGCGTTA TTTATCGGAG TTGCAGTTGC
   GTTAAGAAAT ACGGCCACAA CCCGCGCAAT AAATAGCCTC AACGTCAACG

                                     luc+
                                     =====
351 GCCCGCGAAC GACATTTATA ATGAACGTGA ATTGCTCAAC AGTATGGGCA
   CGGGCGCTTG CTGTAAATAT TACTTGCACT TAACGAGTTG TCATACCCGT

                                     luc+
                                     =====
401 TTTCGCAGCC TACCGTGGTG TTCGTTTCCA AAAAGGGGTT GCAAAAATTT
   AAAGCGTCGG ATGGCACCAC AAGCAAAGGT TTTTCCCCAA CGTTTTTTAA

                                     luc+
                                     =====
451 TTGAACGTGC AAAAAAGCT CCCAATCATC CAAAAATTA TTATCATGGA
   AACTTGCACG TTTTTTTCGA GGGTTAGTAG GTTTTTTAAT AATAGTACCT

                                     luc+
```

Fig. 16 continued

```
=====
501 TTCTAAAACG GATTACCAGG GATTTCAGTC GATGTACACG TTCGTCACAT
   AAGATTTTGC CTAATGGTCC CTAAAGTCAG CTACATGTGC AAGCAGTGTA
                                     luc+
=====
551 CTCATCTACC TCCCGGTTTT AATGAATACG ATTTTGTGCC AGAGTCCTTC
   GAGTAGATGG AGGGCCAAAA TTACTTATGC TAAACACGG TCTCAGGAAG
                                     luc+
=====
601 GATAGGGACA AGACAATTGC ACTGATCATG AACTCCTCTG GATCTACTGG
   CTATCCCTGT TCTGTTAACG TGACTAGTAC TTGAGGAGAC CTAGATGACC
                                     luc+
=====
651 TCTGCCTAAA GGTGTCGCTC TGCCTCATAG AACTGCCTGC GTGAGATTCT
   AGACGGATTT CCACAGCGAG ACGGAGTATC TTGACGGACG CACTCTAAGA
                                     luc+
=====
701 CGCATGCCAG AGATCCTATT TTTGGCAATC AAATCATTCC GGATACTGCG
   GCGTACGGTC TCTAGGATAA AAACCGTTAG TTTAGTAAGG CCTATGACGC
                                     luc+
=====
751 ATTTTAAGTG TTGTTCATT CCATCACGGT TTTGGAATGT TTAATACACT
   TAAATTCAC AACAAGGTAA GGTAAGTCCA AAACCTTACA AATGATGTGA
                                     luc+
=====
801 CGGATATTIG ATATGTGGAT TTCGAGTCGT CTTAATGTAT AGATTTGAAG
   GCCTATAAAC TATACACCTA AAGCTCAGCA GAATTACATA TCTAAACTTC
                                     luc+
=====
851 AAGAGCTGTT TCTGAGGAGC CTTCAGGATT ACAAGATTCA AAGTGCGCTG
   TTCTCGACAA AGACTCCTCG GAAGTCCTAA TGTTCCTAAGT TTCACGCGAC
                                     luc+
=====
901 CTGGTGCCAA CCCTATTCTC CTTCTTCGCC AAAAGCACTC TGATTGACAA
   GACCACGGTT GGGATAAGAG GAAGAAGCGG TTTTCGTGAG ACTAACTGTT
                                     luc+
=====
951 ATACGATTTA TCTAATTAC ACGAAATTGC TTCTGGTGGC GCTCCCCTCT
   TATGCTAAAT AGATTAAATG TGCTTTAACG AAGACCACCG CGAGGGGAGA
                                     luc+
=====
1001 CTAAGGAAGT CGGGGAAGCG GTTGCCAAGA GGTTCATCT GCCAGGTATC
   GATTCCTTCA GCCCCTTCGC CAACGGTTCT CCAAGGTAGA CGGTCCATAG
```

Fig. 16 continued

```
luc+
=====
1051 AGGCAAGGAT ATGGGCTCAC TGAGACTACA TCAGCTATTC TGATTACACC
    TCCGTTCCCTA TACCCGAGTG ACTCTGATGT AGTCGATAAG ACTAATGTGG

luc+
=====
1101 CGAGGGGGAT GATAAACCGG GCGCGGTCGG TAAAGTTGTT CCATTTTGTG
    GCTCCCCCTA CTATTTGGCC CGCGCCAGCC ATTTCAACAA GGTAAAAAAC

luc+
=====
1151 AAGCGAAGGT TGTGGATCTG GATACCGGGA AAACGCTGGG CGTTAATCAA
    TTCGCTTCCA ACACCTAGAC CTATGGCCCT TTTGCGACCC GCAATTAGTT

luc+
=====
1201 AGAGGGCGAAC TGTGTGTGAG AGGTCCTATG ATTATGTCCG GTTATGTAAA
    TCTCCGCTTG ACACACACTC TCCAGGATAC TAATACAGGC CAATACATTT

luc+
=====
1251 CAATCCGGAA GCGACCAACG CCTTGATTGA CAAGGATGGA TGGCTACATT
    GTTAGGCCTT CGCTGGTTGC GGAACAACT GTTCCTACCT ACCGATGTAA

luc+
=====
1301 CTGGAGACAT AGCTTACTGG GACGAAGACG AACACTTCTT CATCGTTGAC
    GACCTCTGTA TCGAATGACC CTGCTTCTGC TTGTGAAGAA GTAGCAACTG

luc+
=====
1351 CGCCTGAAGT CTCTGATTAA GTACAAAGGC TATCAGGTGG CTCCCGCTGA
    GCGGACTTCA GAGACTAATT CATGTTTCCG ATAGTCCACC GAGGGCGACT

luc+
=====
1401 ATTGGAATCC ATCTTGCTCC AACACCCCAA CATCTTCGAC GCAGGTGTGC
    TAACCTTAGG TAGAACGAGG TTGTGGGGTT GTAGAAGCTG CGTCCACAGC

luc+
=====
1451 CAGGTCTTCC CGACGATGAC GCCGGTGAAC TTCCCGCCGC CGTTGTTGTT
    GTCCAGAAGG GCTGCTACTG CGGCCACTTG AAGGGCGGCG GCAACAACAA

luc+
=====
1501 TTGGAGCACG GAAAGACGAT GACGGAAAAA GAGATCGTGG ATTACGTCGC
    AACCTCGTGC CTTTCTGCTA CTGCCTTTTT CTCTAGCACC TAATGCAGCG

luc+
=====
1551 CAGTCAAGTA ACAACCGCGA AAAAGTTGCG CGGAGGAGTT GTGTTTGTGG
    GTCAGTTCAT TGTTGGCGCT TTTTCAACGC GCCTCCTCAA CACAAACACC
```

Fig. 16 continued

```
luc+
=====
1601 ACGAAGTACC GAAAGGTCTT ACCGGAAC TCGACGCAAG AAAAATCAGA
    TGCTTCATGG CTTTCCAGAA TGGCCTTTTG AGCTGCGTTC TTTTGTCT

luc+
=====
1651 GAGATCCTCA TAAAGGCCAA GAAGGGCGGA AAGATCGCCG TGTAATTCTA
    CTCTAGGAGT ATTTCCGGTT CTCCCGCCT TTCTAGCGGC ACATTAAGAT

                                unc-54 3' UTR
                                =====
1701 GGAATTCCAA CTGAGCGCCG GTCGCTACCA TTACCAACTT GTCTGGTGTG
    CCTTAAGGTT GACTCGCGGC CAGCGATGGT AATGGTTGAA CAGACCACAG

                                unc-54 3' UTR
                                =====
1751 AAAAATAATA GGGGCCGCTG TCATCAGAGT AAGTTTAAAC TGAGTTCTAC
    TTTTATTAT CCCCGCGAC AGTAGTCTCA TTCAAATTG ACTCAAGATG

                                unc-54 3' UTR
                                =====
1801 TAACTAACGA GTAATATTTA AATTTTCAGC ATCTCGCGCC CGTGCCTCTG
    ATTGATTGCT CATTATAAAT TTAAAAGTCG TAGAGCGCGG GCACGGAGAC

                                unc-54 3' UTR
                                =====
1851 ACTTCTAAGT CCAATTACTC TTCAACATCC CTACATGCTC TTTCTCCCTG
    TGAAGATTCA GGTAAATGAG AAGTTGTAGG GATGTACGAG AAAGAGGGAC

                                unc-54 3' UTR
                                =====
1901 TGCTCCCACC CCCTATTTTT GTTATTATCA AAAAACTTC TTCTTAATTT
    ACGAGGGTGG GGGATAAAAA CAATAATAGT TTTTTGAAG AAGAATTAAA

                                unc-54 3' UTR
                                =====
1951 CTITGTTTTT TAGCTTCTTT TAAGTCACCT CTAACAATGA AATTGTGTAG
    GAAACAAAAA ATCGAAGAAA ATTCAGTGGA GATTGTTACT TTAACACATC

                                unc-54 3' UTR
                                =====
2001 ATTCAAAAAT AGAATTAATT CGTAATAAAA AGTCGAAAAA AATTGTGCTC
    TAAGTTTTTA TCTTAATTAA GCATTATTTT TCAGCTTTTT TTAACACGAG

                                unc-54 3' UTR
                                =====
2051 CCTCCCCCA TTAATAATAA TTCTATCCA AAATCTACAC AATGTTCTGT
    GGAGGGGGGT AATTATTATT AAGATAGGGT TTAGATGTG TTACAAGACA

                                unc-54 3' UTR
                                =====
2101 GTACACTTCT TATGTTTTTT TTAATCTCTGA TAAATTTTTT TTGAAACATC
```

Fig. 16 continued

```
CATGTGAAGA ATACAAAAA AATGAAGACT ATTTAAAAA AACTTTGTAG
unc-54 3' UTR
=====
2151 ATAGAAAAA CCGCACACAA AATACCTTAT CATATGTTAC GTTTCAGTTT
TATCTTTTTT GCGTGTGTT TTATGGAATA GTATACAATG CAAAGTCAAA
unc-54 3' UTR
=====
2201 ATGACCGCAA TTTTATTTC TTCGCACGTC TGGGCCTCTC ATGACGTCAA
TACTGGCGTT AAAAATAAAG AAGCGTGCAG ACCCGGAGAG TACTGCAGTT
unc-54 3' UTR
=====
2251 ATCATGCTCA TCGTGAAAAA GTTTTGGAGT ATTTTGGAA TTTTCAATC
TAGTACGAGT AGCACTTTT CAAAACCTCA TAAAACCTT AAAAAGTTAG
unc-54 3' UTR
=====
2301 AAGTGAAAGT TTATGAAATT AATTTTCCTG CTTTGTCTT TTGGGGGTTT
TTCACTTTCA AATACTTTAA TTAAAGGAC GAAAACGAAA AACCCCCAAA
unc-54 3' UTR
=====
2351 CCCCTATTGT TTGTCAAGAG TTTCGAGGAC GCGTTTTTC TTGCTAAAT
GGGGATAACA AACAGTTCTC AAAGCTCCTG CCGCAAAAAG AACGATTTTA
unc-54 3' UTR
=====
2401 CACAAGTATT GATGAGCAGC ATGCAAGAAA GATCGGAAGA AGGTTTGGGT
GTGTTCAATA CTACTCGTGC TACGTTCTTT CTAGCCTTCT TCCAACCCA
unc-54 3' UTR
=====
2451 TTGAGGCTCA GTGGAAGGTG AGTAGAAGTT GATAATTTGA AAGTGGAGTA
AACTCCGAGT CACCTTCCAC TCATCTTCAA CTATTAACT TTCACCTCAT
unc-54 3' UTR
=====
2501 GTGTCATGG GGTTTTGGCC TTAAATGACA GAATACATTC CCAATATACC
CACAGATACC CAAAAACGG AATTACTGT CTTATGTAAG GGTATATGG
unc-54 3' UTR
=====
2551 AAACATAACT GTTTCCTACT AGTCGGCCGT ACGGGCCCTT TCGTCTCGCG
TTTGTATTGA CAAAGGATGA TCAGCCGGCA TGCCCGGGA AGCAGAGCGC
2601 CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC TCCCGGAGAC
GCAAAGCCAC TACTGCCACT TTTGGAGACT GTGTACGTCG AGGGCCTCTG
2651 GGTACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG
CCAGTGTGCA ACAGACATTC GCCTACGGCC CTCGTCTGTT CGGGCAGTCC
2701 GCGCGTCAGC GGGTGTGGC GGGTGTGGG GCTGGCTTAA CTATGCGGCA
```

Fig. 16 continued

```
CGCGCAGTCG CCCACAACCG CCCACAGCCC CGACCGAATT GATACGCCGT
2751 TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGBA
      AGTCTCGTCT AACATGACTC TCACGTGGTA TACGCCACAC TTTATGGCGT
2801 CAGATGCGTA AGGAGAAAAT ACCGCATCAG GCGGCCTTAA GGGCCTCGTG
      GTCTACGCAT TCCTCTTTTA TGGCGTAGTC CGCCGGAATT CCCGGAGCAC
2851 ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC
      TATGCGGATA AAAATATCCA ATTACAGTAC TATTATTACC AAAGAATCTG
2901 GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT
      CAGTCCACCG TGAAAAGCCC CTTTACACGC GCCTTGGGGA TAAACAAATA
2951 TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA
      AAAAGATTTA TGTAAGTTTA TACATAGGCG AGTACTCTGT TATTGGGACT

                                          amp
                                          =====
3001 TAAATGCTTC AATAATATTG AAAAAGGAAG AGTATGAGTA TTCAACATTT
      ATTTACGAAG TTATTATAAC TTTTTCCTTC TCATACTCAT AAGTTGTAAA

                                          amp
                                          =====
3051 CCGTGTGCGC CTTATTCCCT TTTTTCGGGC ATTTTGCCTT CCGTGTTCCTG
      GGCACAGCGG GAATAAGGGA AAAAACGCGG TAAAACGGAA GGACAAAAAC

                                          amp
                                          =====
3101 CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT
      GAGTGGGTCT TTGCGACCAC TTTCATTTTC TACGACTTCT AGTCAACCCA

                                          amp
                                          =====
3151 GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA
      CGTGCTCACC CAATGTAGCT TGACCTAGAG TTGTCGCCAT TCTAGGAAC

                                          amp
                                          =====
3201 GAGTTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC
      CTCAAAGCG GGGCTTCTTG CAAAAGGTTA CTAAGCTGTA AAATTTCAAG

                                          amp
                                          =====
3251 TGCTATGTGG CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC
      ACGATACACC GCGCCATAAT AGGGCATAAC TGCGGCCCGT TCTCGTTGAG

                                          amp
                                          =====
3301 GGTGCGCGCA TACACTATTC TCAGAAATGAC TTGGTTGAGT ACTCACCAGT
      CCAGCGGCGT ATGTGATAAG AGTCTTACTG AACCAACTCA TGAGTGGTCA

                                          amp
                                          =====
```

Fig. 16 continued

3351 CACAGAAAAG CATCTTACGG ATGGCATGAC ACTAAGAGAA TTATGCAGTG  
GTGTCTTTTC GTAGAATGCC TACCGTACTG TCATTCTCTT AATACGTCAC  
amp  
=====

3401 CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG  
GACGGTATTG GTACTCACTA TTGTGACGCC GGTGAATGA AGACTGTGTC  
amp  
=====

3451 ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TGCACAACA TGGGGGATCA  
TAGCCTCCTG GCTTCCTCGA TTGGCGAAAA AACGTGTTGT ACCCCCTAGT  
amp  
=====

3501 TGTAATCGC CTTGATCGTT GGAACCGGA GCTGAATGAA GCCATACCAA  
ACATTGAGCG GAACTAGCAA CCCTTGGCCT CGACTTACTT CGGTATGTT  
amp  
=====

3551 ACGACGAGCG TGACACCACG ATGCCTGTAG CAATGGCAAC AACGTTGCGC  
TGCTGCTCGC ACTGTGGTGC TACGGACATC GTTACCGTTG TTGCAACGCG  
amp  
=====

3601 AACTATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAAATTAAT  
TTTGATAATT GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTTAATTA  
amp  
=====

3651 AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCCC  
TCTGACCTAC CTCCGCCTAT TTCAACGTCC TGGTGAAGAC GCGAGCCGGG  
amp  
=====

3701 TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG  
AAGGCCGACC GACCAAATAA CGACTATTTA GACCTCGGCC ACTCGCACCC  
amp  
=====

3751 TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT  
AGAGCGCCAT AGTAACGTCG TGACCCCGGT CTACCATTCTG GGAGGGCATA  
amp  
=====

3801 CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA  
GCATCAATAG ATGTGCTGCC CCTCAGTCCG TTGATACCTA CTTGCTTTAT  
amp  
=====

3851 GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACGTCA  
CTGTCTAGCG ACTCTATCCA CGGAGTGACT AATTCGTAA CATTGACAGT  
3901 GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC TTCATTTTAA

Fig. 16 continued

CTGGTTCAAA TGAGTATATA TGAAATCTAA CTAAATTTTG AAGTAAAAAT

3951 ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA  
TAAATTTTCC TAGATCCACT TCTAGGAAAA ACTATTAGAG TACTGGTTTT

4001 TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAAG  
AGGGAATTGC ACTCAAAAGC AAGGTGACTC GCAGTCTGGG GCATCTTTTC

4051 ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT  
TAGTTTCCTA GAAGAACTCT AGGAAAAAAA GACGCGCA'T' AGACGACGAA

4101 GCAAACAAAA AAACCACCGC TACCAGCGGT GGTTCGTTTG CCGGATCAAG  
CGTTTGTTTT TTTGGTGGCG ATGGTCGCCA CCAAACAAAC GGCCTAGTTC

4151 AGCTACCAAC TCTTTTTCCG AAGGTAAGT GCTTCAGCAG AGCGCAGATA  
TCGATGTTG AGAAAAAGGC TTCCATTGAC CGAAGTCGTC TCGCGTCTAT

4201 CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA  
GGTTTATGAC AGGAAGATCA CATCGGCATC AATCCGGTGG TGAAGTTCTT

4251 CTCTGTAGCA CCGCTACAT ACCTCGCTCT GCTAATCCTG TTACCAGTGG  
GAGACATCGT GCGGGATGTA TGGAGCGAGA CGATTAGGAC AATGGTCACC

4301 CTGCTGCCAG TGGCGATAAG TCGTGCTTGA CCGGGTTGGA CTCAAGACGA  
GACGACGGTC ACCGCTATTC AGCACAGAAT GGCCCAACCT GAGTTCTGCT

4351 TAGTTACCGG ATAAGGCGCA GCGGTGCGGC TGAACGGGGG GTTCGTGCAC  
ATCAATGGCC TATTCCGCGT CGCCAGCCCG ACTTGCCCCC CAAGCACGTG

4401 ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC  
TGTCGGGTCG AACCTCGCTT GCTGGATGTG GCTTGACTCT ATGGATGTG

4451 GTGAGCATTG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GGCGGACAGG  
CACTCGTAAC TCTTTCGCGG TGCGAAGGGC TTCCCTCTTT CCGCCTGTCC

4501 TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC  
ATAGGCCATT CGCCGTCCCA GCCTTGTCTT CTCGCGTGCT CCCTCGAAGG

4551 AGGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT  
TCCCCCTTTG CGGACCATAG AAATATCAGG ACAGCCCAA GCGGTGGAGA

4601 GACTTGAGCG TCGATTTTTC TGATGCTCGT CAGGGGGGCG GAGCCTATGG  
CTGAACTCGC AGCTAAAAAC ACTACGAGCA GTCCCCCGC CTCGGATACC

4651 AAAACGCCA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC  
TTTTTGCGGT CGTTGCGCCG GAAAAATGCC AAGGACCGGA AAACGACCGG

4701 TTTTGCTCAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC  
AAAACGAGTG TACAAGAAAG GACGCAATAG GGGACTAAGA CACCTATTGG

4751 GTATTACCGC CTTTGAGTGA GCTGATACCG CTCGCCGAG CCGAACGACC  
CATAATGGCG GAAACTCACT CGACTATGGC GAGCGGCGTC GGCTTGCTGG

4801 GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA

Fig. 16 continued

```

CTCGCGTCGC TCAGTCACTC GCTCCTTCGC CTTCTCGCGG GTTATGCGTT
4851 ACCGCCTCTC CCCGCGCGTT GGCCGATTCA TTAATGCAGC TGGCACGACA
TGGCGGAGAG GGGCGCGCAA CCGGCTAAGT AATTACGTCG ACCGTGCTGT
4901 GGTTCCTCCGA CTGGAAAGCG GGCAGTGAGC GCAACGCAAT TAATGTGAGT
CCAAAGGGCT GACCTTTCGC CCGTCACTCG CGTTGCGTTA ATTACACTCA
4951 TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC TTCCGGCTCG
ATCGAGTGAG TAATCCGTGG GGTCCGAAAT GTGAAATACG AAGGCCGAGC
5001 TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTACACA GGAACAGCT
ATACAACACA CCTTAACACT CGCCTATTGT TAAAGTGTGT CCTTGTCGA
5051 ATGACCATGA TTACGCCAAG CTGTAAGTTT AAACATGATC TTACTAACTA
TACTGGTACT AATGCGGTTC GACATTCAAA TTTGTACTAG AATGATTGAT
5101 ACTATTCTCA TTTAAATTTT CAGAGCTTAA AAATGGCTGA AATCACTCAC
TGATAAGAGT AAATTTAAAA GTCTCGAATT TTTACCGACT TTAGTGAGTG
5151 AACGATGGAT ACGCTAACAA CTTGGAAATG AAATAAGCTT GCATGCCTGC
TTGCTACCTA TGCGATTGTT GAACCTTTAC TTTATTGAA CGTACGGACG

vit-2 promoter
=====
StuI
~~~~~
5201 AGGCCTTGGT CGACTCTAGA GGATCAAAC TATTACTTG AAACAATTTA
TCCGGAACCA GCTGAGATCT CCTAGTTTGA CATAATGAAC TTTGTTAAAT

vit-2 promoter
=====
5251 GTTATATGTT TAGAACCCCT CATTCAAAAT TAATAGACAG GGCTCTCACC
CAATATACAA ATCTTGGGGA GTAAGTTTTA ATTATCTGTC CCGAGAGTGG

vit-2 promoter
=====
5301 GAATGTTGCA ATTTGTTTCT GATAAGGGTC ACAAAGCGGA GCGAATGCTT
CTTACAACGT TAAACAAAGA CTATTCCCAG TGTTCGCCT CGCTTACGAA

vit-2 promoter
=====
5351 GAATGTGTCC ATCAATGAGC TTATCAATGC GCTAAAACGC TATAACTTCC
CTTACACAGG TAGTTACTCG AATAGTTACG CGATTTTGCG ATATTGAAGG

vit-2 promoter
=====
5401 ATATGAAGTC AATCGAACAT ATGTCAATCT TTAGCCSTAT ATAAAGGTGC
TATACTTCAG TTAGCTTGTA TACAGTTAGA AATCGGCATA TATTTCCACG

vit-2 promoter exon 1 (in frame - partial)
=====
5451 ACTGAAAACA GTCCAATCAC GGTTCAGCCA TGAGGTCGAT CCCCAGCCGG
TGACTTTTGT CAGGTTAGTG CCAAGTCGGT ACTCCAGCTA GGGGCCGGCC

```

Fig. 16 continued

```
exon 1 (in frame - partial) synth. intron
=====
5501 GATTGGCCAA AGGACCCAAA GGTATGTTTC GAATGATACT AACATAACAT
 CTAACCGGTT TCCTGGGTTT CCATACAAAG CTTACTATGA TTGTATTGTA

synth. intron
=====
5551 AGAACATTTT CAGGAGGACC CTTGGAGGGT ACCGGGGATT GGCCAAAGGA
 TCTTGTAATAA GTCCTCCTGG GAACCTCCCA TGGCCCCTAA CCGGTTTCCT

5601 CCCAAAGGTA TGTTTCGAAT GATACTAACA TAACATAGAA CATTTCAGG
 GGGTTTCCAT ACAAAGCTTA CTATGATTGT ATTGTATCTT GTAAAAGTCC

 SacI
                                ~~~~~
5651 AGGACCCTTG CTTGGAGGGT ACCGAGCTCA GAAAAA
    TCCTGGGAAC GAACCTCCCA TGGCTCGAGT CTTTTT
```

fig. 17

## III. Predicted DNA sequence pGQ2

```

                                     NLS                                     luc+
=====
1  ATGACTGCTC CAAAGAAGAA GCGTAAGGTA CCGGTAGAAA AAATGGAAGA
   TACTGACGAG GTTCTTCTT CGCATTCAT  GGCCATCTTT TTTACCTTCT

                                     luc+
=====
51 CGCCAAAAAC ATAAAGAAAG GCCCGGCGCC ATTCTATCCG CTGGAAGATG
   GCGGTTTTTG TATTTCTTC CGGGCCGCGG TAAGATAGGC GACCTTCIAC

                                     luc+
=====
101 GAACCGCTGG AGAGCAACTG CATAAGGCTA TGAAGAGATA CGCCCTGCTT
   CTGGCGACC TCTCGTTGAC GTATTCCGAT ACTTCTCTAT GCGGGACCAA

                                     luc+
=====
151 CCTGGAACAA TTGCTTTTAC AGATGCACAT ATCGAGGTGG ACATCACTTA
   GGACCTTGTT AACGAAAATG TCTACGTGTA TAGCTCCACC TGTAGTGAAT

                                     luc+
=====
201 CGCTGAGTAC TTCGAAATGT CCGTTCGGTT GGCAGAAGCT ATGAAACGAT
   GCGACTCATG AAGCTTTACA GGCAAGCCAA CCGTCTTCGA TACTTTGCTA

                                     luc+
=====
251 ATGGGCTGAA TACAAATCAC AGAATCGTCG TATGCAGTGA AAACCTCTCTT
   TACCCGACTT ATGTTTAGTG TCTTAGCAGC ATACGTCACI TTTGAGAGAA

                                     luc+
=====
301 CAATTCTTTA TGCCGGTGTT GGGCGCGTTA TTTATCGGAG TTGCAGTTGC
   GTTAAGAAAT ACGGCCACAA CCCGCGCAAT AAATAGCCTC AACGTCAACG

                                     luc+
=====
351 GCCCGCGAAC GACATTTATA ATGAACGTGA ATTGCTCAAC AGTATGGGCA
   CGGCGCCTTG CTGTAAATAT TACTTGCACT TAACGAGTTG TCATACCCGT

                                     luc+
=====
401 TTTCGCAGCC TACCGTGGTG TTCGTTTCCA AAAAGGGGTT GCAAAAAATT
   AAAGCGTCGG ATGGCACCAC AAGCAAAGGT TTTTCCCCAA CGTTTTTTAA

                                     luc+
=====
451 TTGAACGTGC AAAAAAAGCT CCCAATCATC CAAAAAATTA TTATCATGGA
   AACTTGACG TTTTTTTCGA GGGTTAGTAG GTTTTTTAAT AATAGTACCT

```

Fig. 17 continued

```

=====
luc+
=====
501 TTCTAAAACG GATTACCAGG GATTTCAGTC GATGTACACG TTCGTACAT
    AAGATTTTGC CTAATGGTCC CTAAAGTCAG CTACATGTGC AAGCAGTGTA

=====
luc+
=====
551 CTCATCTACC TCCCGGTTTT AATGAATACG ATTTTGTGCC AGAGTCCTTC
    GAGTAGATGG AGGGCCAAAA TTAATTATGC TAAACACGG TCTCAGGAAG

=====
luc+
=====
601 GATAGGGACA AGACAATTGC ACTGATCATG AACTCCTCTG GATCTACTGG
    CTATCCCTGT TCTGTTAACG TGAAGTAGTAC TTGAGGAGAC CTAGATGACC

=====
luc+
=====
651 TCTGCCTAAA GGTGTCGCTC TGCCTCATAG AACTGCCTGC GTGAGATTCT
    AGACGGATTT CCACAGCGAG ACGGAGTATC TTGACGGACG CACTCTAAGA

=====
luc+
=====
701 CGCATGCCAG AGATCCTATT TTTGGCAATC AAATCATTCC GGATACTGCG
    GCGTACGGTC TCTAGGATAA AAACCGTTAG TTTAGTAAGG CCTATGACGC

=====
luc+
=====
751 ATTTTAAGTG TTGTTCCATT CCATCACGGT TTTGGAATGT TTAATACACT
    TAAAATTCAC AACAAGGTAA GGTAGTGCCA AAACCTTACA AATGATGTGA

=====
luc+
=====
801 CCGATATTTG ATATGIGGAT TTCGAGTCGT CTTAATGTAT AGATTTGAAG
    GCCTATAAAC TATACACCTA AAGCTCAGCA GAATTACATA TCTAACTTC

=====
luc+
=====
851 AAGAGCTGTT TCTGAGGAGC CTTCAGGATT ACAAGATTCA AAGTGCGCTG
    TTCTCGACAA AGACTCCTCG GAAGTCCTAA TGTTCTAAGT TTCACGCGAC

=====
luc+
=====
901 CTGGTGCCAA CCCTATTCTC CTTCTTCGCC AAAAGCACTC TGATTGACAA
    GACCACGGTT GGGATAAGAG GAAGAAGCGG TTTTCGTGAG ACTAAGTGT

=====
luc+
=====
951 ATACGATTTA TCTAATTTAC ACGAAATTGC TTCTGGTGGC GCTCCCTCT
    TATGCTAAAT AGATTAAATG TGCTTTAACG AAGACCACCG CGAGGGGAGA

=====
luc+
=====
1001 CTAAGGAAGT CGGGGAAGCG GTTGCCAAGA GGTTCCATCT GCCAGGTATC
    GATTCCTTCA GCCCCTTCGC CAACGGTTCT CCAAGGTAGA CGGTCCATAG
=====
```

Fig. 17 continued

```
luc+
=====
1051 AGGCAAGGAT ATGGGCTCAC TGAGACTACA TCAGCTATTC TGATTACACC
    TCCGTTCCCTA TACCCGAGTG ACTCTGATGT AGTCGATAAG ACTAATGTGG

luc+
=====
1101 CGAGGGGGAT GATAAACCGG GCGCGGTCGG TAAAGTTGTT CCATTTTTTG
    GCTCCCCCTA CTATTTGGCC CGCGCCAGCC ATTTCAACAA GGTA AAAAAC

luc+
=====
1151 AAGCGAAGGT TGTGGATCTG GATACCGGGA AAACGCTGGG CGTTAATCAA
    TTCGCTTCCA ACACCTAGAC CTATGGCCCT TTGCGACCC GCAATTAGTT

luc+
=====
1201 AGAGGCGAAC TGTGTGTGAG AGGTCCTATG ATTATGTCCG GTTATGTAAA
    TCTCCGCTTG ACACACACTC TCCAGGATAC TAATACAGGC CAATACATTT

luc+
=====
1251 CAA TCCGGAA GCGACCAACG CCTTGATTGA CAAGGATGGA TGGCTACATT
    GTTAGGCCTT CGCTGGTTGC GGAAC TAACT GTTCCTACCT ACCGATGTA

luc+
=====
1301 CTGGAGACAT AGCTTACTGG GACGAAGACG AACACTTCTT CATCGTTGAC
    GACCTCTGTA TCGAATGACC CTGCTTCTGC TTGTGAAGAA GTAGCAACTG

luc+
=====
1351 CGCCTGAAGT CTCTGATTAA GTACAAAGGC TATCAGGTGG CTCCCGCTGA
    GCGGACTTCA GAGACTAATT CATGTTTCCG ATAGTCCACC GAGGGCGACT

luc+
=====
1401 ATTGGAATCC ATCTTGCTCC AACACCCCAA CATCTTCGAC GCAGGTGTGC
    TAACCTTAGG TAGAACGAGG TTGTGGGGTT GTAGAAGCTG CGTCCACAGC

luc+
=====
1451 CAGGTCTTCC CGACGATGAC GCCGGTGAAC TTCCCGCCGC CGTTGTTGTT
    GTCCAGAAGG GCTGCTACTG CGGCCACTTG AAGGGCGGCG GCAACAACAA

luc+
=====
1501 TTGGAGCACG GAAAGACGAT GACGGAAAAA GAGATCGTGG ATTACGTCGC
    AACCTCGTGC CTTTCTGCTA CTGCCTTTT CTCTAGCACC TAATGCAGCG

luc+
=====
1551 CAGTCAAGTA ACAACCGCGA AAAAGTTGCG CGGAGGAGTT GTGTTTGTGG
```

Fig. 17 continued

```
GTCAGTTCAT TGTTGGCGCT TTTTCAACGC GCCTCCTCAA CACAAACACC
luc+
=====
1601 ACGAAGTACC GAAAGGTCTT ACCGGAAAAC TCGACGCAAG AAAAATCAGA
    TGCTTCATGG CTTTCCAGAA TGGCCTTTTG AGCTGCGTTC TTTTtagTCT
luc+
=====
1651 GAGATCCTCA TAAAGGCCAA GAAGGGCGGA AAGATCGCCG TGTAATTCTA
    CTCTAGGAGT ATTTCCGGTT CTCCCGCCT TTCTAGCGGC ACATTAAGAT
                                unc-54 3' UTR
                                =====
1701 GGAATTCCAA CTGAGCGCCG GTCGCTACCA TTACCAACTT GTCTGGTGTC
    CCTTAAGGTT GACTCGCGGC CAGCGATGGT AATGGTTGAA CAGACCAGAG
                                unc-54 3' UTR
                                =====
1751 AAAAATAATA GGGGCCGCTG TCATCAGAGT AAGTTTAAAC TGAGTTCTAC
    TTTTtATTAT CCCCGGCGAC AGTAGTCTCA TTCAAATTG ACTCAAGATG
                                unc-54 3' UTR
                                =====
1801 TAACTAACGA GTAATATTTA AATTTCAGC ATCTCGCGCC CGTGCCTCTG
    ATTGATTGCT CATTATAAAT TTAAAAGTCG TAGAGCGCGG GCACGGAGAC
                                unc-54 3' UTR
                                =====
1851 ACTTCTAAGT CCAATTACTC TTCAACATCC CTACATGCTC TTTCTCCCTG
    TGAAGATTCA GGTTAATGAG AAGTTGTAGG GATGTACGAG AAAGAGGGAC
                                unc-54 3' UTR
                                =====
1901 TGCTCCCACC CCCTATTTTT GTTATTATCA AAAAAACTTC TTCTTAATTT
    ACGAGGGTGG GGGATAAAAA CAATAATAGT TTTTTTGAAG AAGAATTAAA
                                unc-54 3' UTR
                                =====
1951 CTTTGTTTTT TAGCTTCTTT TAAGTCACCT CTAACAATGA AATTGTGTAG
    GAAACAAAAA ATCGAAGAAA ATTCAGTGGA GATTGTTACT TTAACACATC
                                unc-54 3' UTR
                                =====
2001 ATTCAAAAAT AGAATTAATT CGTAATAAAA AGTCGAAAAA AATTGTGCTC
    TAAGTTTTTA TCTTAATTAA GCATTATTTT TCAGCTTTTT TTAACACGAG
                                unc-54 3' UTR
                                =====
2051 CCTCCCCCA TTAATAATAA TTCTATCCA AAATCTACAC AATGTTCTGT
    GGAGGGGGGT AATTATTATT AAGATAGGGT TTTAGATGTG TTACAAGACA
                                unc-54 3' UTR
                                =====
```

Fig. 17 continued

```

2101 GTACACTTCT TATGTTTTTT TTA CTCTCTGA TAAATTTTTT TTGAAACATC
    CATGTGAAGA ATACAAAAAA AATGAAGACT ATTTAAAAAA AACTTTGTAG

unc-54 3' UTR
=====
2151 ATAGAAAAAA CCGCACACAA AATACCTTAT CATATGTTAC GTTTCAGTTT
    TATCTTTTTT GCGTGTGTT TTATGGAATA GTATACAATG CAAAGTCAAA

unc-54 3' UTR
=====
2201 ATGACCGCAA TTTTATTTT TC GCACGTC TGGGCCTCTC ATGACGTCAA
    TACTGGCGTT AAAAATAAAG AAGCGTGCAG ACCCGGAGAG TACTGCAGTT

unc-54 3' UTR
=====
2251 ATCATGCTCA TCGTGAAAAA GTTTTGGAGT ATTTTGGAA TTTTCAATC
    TAGTACGAGT AGCACTTTTT CAAAACCTCA TAAAAACCTT AAAAAGTTAG

unc-54 3' UTR
=====
2301 AAGTGAAAGT TTATGAAATT AATTTTCCTG CTTTGTCTT TTGGGGGTTT
    TTCACTTTCA AATACTTTAA TTAAAAGGAC GAAAACGAAA AACCCCCAAA

unc-54 3' UTR
=====
2351 CCCCTATTGT TTGTCAAGAG TTTCGAGGAC GCGTTTTTTC TTGCTAAAT
    GGGGATAACA AACAGTTCTC AAAGCTCCTG CCGCAAAAAG AACGATTTTA

unc-54 3' UTR
=====
2401 CACAAGTATT GATGAGCAGC ATGCAAGAAA GATCGGAAGA AGGTTTGGGT
    GTGTTTCAAA CTACTCGTGC TACGTTCTTT CTAGCCTTCT TCCAAACCCA

unc-54 3' UTR
=====
2451 TTGAGGCTCA GTGGAAGGTG AGTAGAAGTT GATAATTGA AAGTGGAGTA
    AACTCCGAGT CACCTTCCAC TCATCTTCAA CTATTAACT TTCACCTCAT

unc-54 3' UTR
=====
2501 GTGTCTATGG GTTTTTTGCC TTAAATGACA GAATACATTC CCAATATACC
    CACAGATACC CCAAAAACGG AATTTACTGT CTTATGTAAG GGTATATGG

unc-54 3' UTR
=====
2551 AAACATAACT GTTTCCTACT AGTCGGCCGT ACGGGCCCTT TCGTCTCGCG
    TTTGTATTGA CAAAGGATGA TCAGCCGGCA TGCCCGGGA AGCAGAGCGC

2601 CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC TCCCGGAGAC
    GCAAAGCCAC TACTGCCACT TTTGGAGACT GTGTACGTCG AGGGCCTCTG

2651 GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG
    CCAGTGTCGA ACAGACATTC GCCTACGGCC CTCGTCTGTT CGGGCAGTCC

```

MSC II

## Fig. 17 continued

2701 GCGCGTCAGC GGGTGTGGC GGGTGTGGG GCTGGCTTAA CTATGCGGCA  
CGCGCAGTCG CCCACAACCG CCCACAGCCC CGACCGAATT GATACGCCGT

2751 TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA  
AGTCTCGTCT AACATGACTC TCACGTGGTA TACGCCACAC TTTATGGCGT

2801 CAGATGCGTA AGGAGAAAAT ACCGCATCAG GCGGCCTTAA GGGCCTCGTG  
GTCTACGCAT TCCTCTTTTA TGGCGTAGTC CGCCGGAATT CCCGAGCAC

2851 ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC  
TATGCGGATA AAAATATCCA ATTACAGTAC TATTATTACC AAAGAATCTG

2901 GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT  
CAGTCCACCG TGAAAAGCCC CTTTACACGC GCCTTGGGGA TAAACAAATA

2951 TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA  
AAAAGATTTA TGTAAGTTTA TACATAGGCG AGTACTCTGT TATTGGGACT

amp  
=====

3001 TAAATGCTTC AATAATATTG AAAAAGGAAG AGTATGAGTA TTCAACATTT  
ATTTACGAAG TTATTATAAC TTTTTCCTTC TCATACTCAT AAGTTGTAA

amp  
=====

3051 CCGTGTGCGC CTTATTCCTT TTTTTCGGC ATTTTGCCTT CCTGTTTTTG  
GGCACAGCGG GAATAAGGGA AAAAACGCGG TAAACGGAA GGACAAAAC

amp  
=====

3101 CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT  
GAGTGGGTCT TTGCGACCAC TTTCATTTTC TACGACTTCT AGTCAACCCA

amp  
=====

3151 GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA  
CGTGCTCACC CAATGTAGCT TGACCTAGAG TTGTCGCCAT TCTAGGAAT

amp  
=====

3201 GAGTTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTAAAGTTC  
CTCAAAGCG GGGCTTCTTG CAAAAGGTTA CTAATCGTGA AAATTTCAG

amp  
=====

3251 TGCTATGTGG CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC  
ACGATACACC GCGCCATAAT AGGGCATAAC TGCGGCCCGT TCTCGTTGAG

amp  
=====

3301 GGTGCGCGCA TACACTATTC TCAGAATGAC TTGGTTGAGT ACTCACCAGT  
CCAGCGGCGT ATGTGATAAG AGTCTTACTG AACCAACTCA TGAGTGGTCA

amp

Fig. 17 continued

=====  
3351 CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG  
GTGTCTTTTC GTAGAATGCC TACCGTACTG TCATTCTCTT AATACGTCAC

amp

=====  
3401 CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG  
GACGGTATTG GTACTCACTA TTGTGACGCC GGTGAATGA AGACTGTTGC

amp

=====  
3451 ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA  
TAGCCTCCTG GCTTCCTCGA TTGGCGAAAA AACGTGTTGT ACCCCCTAGT

amp

=====  
3501 TGTAACTCGC CTTGATCGTT GGAACCGGA GCTGAATGAA GCCATACCAA  
ACATTGAGCG GAACTAGCAA CCCTTGGCCT CGACTTACTT CGGTATGGTT

amp

=====  
3551 ACGACGAGCG TGACACCACG ATGCCTGTAG CAATGGCAAC AACGTTGCGC  
TGCTGCTCGC ACTGTGGTGC TACGGACATC GTTACCGTTG TTGCAACGCG

amp

=====  
3601 AAACATATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAAATTAAT  
TTTGATAATT GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTTAATTA

amp

=====  
3651 AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCC  
TCTGACCTAC CTCCGCCTAT TTCAACGTCC TGGTGAAGAC GCGAGCCGGG

amp

=====  
3701 TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG  
AAGGCCGACC GACCAAATAA CGACTATTTA GACCTCGGCC ACTCGCACCC

amp

=====  
3751 TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT  
AGAGCGCCAT AGTAACGTCG TGACCCCGGT CTACCATTCTG GGAGGGCATA

amp

=====  
3801 CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA  
GCATCAATAG ATGTGCTGCC CCTCAGTCCG TTGATACCTA CTGCTTTAT

amp

=====  
3851 GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACGTCA  
CTGTCTAGCG ACTCTATCCA CGGAGTGACT AATTCGTAAC CATTGACAGT

Fig. 17 continued

```

3901  GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC TTCATTTTTA
      CTGGTTCAAA TGAGTATATA TGAAATCTAA CTAAATTTTG AAGTAAAAAT

3951  ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA
      TAAATTTTCC TAGATCCACT TCTAGGAAAA ACTATTAGAG TACTGGTTTT

4001  TCCCTTAACG TGAGTTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG
      AGGGAATTGC ACTCAAAGC AAGGTGACTC GCAGTCTGGG GCATCTTTTC

4051  ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT
      TAGTTTCCTA GAAGAACTCT AGGAAAAAAA GACGCGCATT AGACGACGAA

4101  GCAAACAAAA AAACCACCGC TACCAGCGGT GGTGTGTTTG CCGGATCAAG
      CGTTTGTTTT TTTGGTGGCG ATGGTCGCCA CCAAACAAAC GGCCTAGTTC

4151  AGCTACCAAC TCTTTTTCGG AAGGTAAC TGCTTCAGCAG AGCGCAGATA
      TCGATGGTTG AGAAAAAGGC TTCCATTGAC CGAAGTCGTC TCGCGTCTAT

4201  CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA
      GGTTCATGAC AGGAAGATCA CATCGGCATC AATCCGGTGG TGAAGTTCTT

4251  CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACCAGTGG
      GAGACATCGT GCGGGATGTA TGGAGCGAGA CGATTAGGAC AATGGTCACC

4301  CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA
      GACGACGGTC ACCGCTATTC AGCACAGAAT GGCCCAACCT GAGTTCTGCT

4351  TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC
      ATCAATGGCC TATTCGCGT CGCCAGCCCG ACTTGCCCCC CAAGCACGTG

4401  ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC
      TGTCGGGTCG AACCTCGCTT GCTGGATGTG GCTTGACTCT ATGGATGTGC

4451  GTGAGCATTG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GCGCGACAGG
      CACTCGTAAC TCTTTCGCGG TGCGAAGGGC TTCCCTCTTT CCGCCTGTCC

4501  TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC
      ATAGGCCATT CGCCGTCCCA GCCTTGTCTT CTCGCGTGCT CCCTCGAAGG

4551  AGGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT
      TCCCCCTTTG CGGACCATAG AAATATCAGG ACAGCCCAA GCGGTGGAGA

4601  GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG GAGCCTATGG
      CTGAACTCGC AGCTAAAAAC ACTACGAGCA GTCCCCCGC CTCGGATACC

4651  AAAACGCECA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC
      TTTTTCGGT CGTTGCGCG GAAAAATGCC AAGGACCGGA AAACGACCGG

4701  TTTTGCTCAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC
      AAAACGAGTG TACAAGAAAG GACGCAATAG GGGACTAAGA CACCTATTGG

4751  GTATTACCGC CTTTGAGTGA GCTGATACCG CTCGCCGAG CCGAACGACC
      CATAATGGCG GAAACTCACT CGACTATGGC GAGCGGCGTC GGCTTGCTGG

```

Fig. 17 Continued

4801 GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA  
CTCGCGTCGC TCAGTCACTC GTCCTTCGC CTTCTCGCGG GTTATGCGTT

4851 ACCGCCTCTC CCCGCGCGTT GGCCGATTCA TTAATGCAGC TGGCAGCACA  
TGGCGGAGAG GGGCGCGCAA CCGGCTAAGT AATTACGTCG ACCGTGCTGT

4901 GGTTCCTCCGA CTGGAAAGCG GGCAGTGAGC GCAACGCAAT TAATGTGAGT  
CCAAAGGGCT GACCTTTCGC CCGTCACTCG CGTTGCGTTA ATTACACTCA

4951 TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC TTCCGGCTCG  
ATCGAGTGAG TAATCCGTGG GGTCCGAAAT GTGAAATACG AAGGCCGAGC

5001 TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTACACACA GGAAACAGCT  
ATACAACACA CCTTAACACT CGCCTATTGT TAAAGTGTGT CCTTTGTGCA

5051 ATGACCATGA TTACGCCAAG CTGTAAGTTT AAACATGATC TTAATAACTA  
TACTGGTACT AATGCGGTTC GACATTCAAA TTTGTACTAG AATGATTGAT

5101 ACTATTCTCA TTTAAATTTT CAGAGCTTAA AAATGGCTGA AATCACTCAC  
TGATAAGAGT AAATTTAAAA GTCTCGAATT TTTACCGACT TTAGTGAGTG

5151 AACGATGGAT ACGCTAACAA CTTGGAAATG AAATAAGCTT GCATGCCTGC  
TTGCTACCTA TGCGATTGTT GAACCTTTAC TTTATTTCGAA CGTACGGAGC

ctl-1 promoter + coding region

=====

o-GQ3

=====

StuI

5201 AGGCCTGAGA TATTTTGC GC TCAATATG TTTTGTGTCC CCGTAATATT  
TCCGGACTCT ATAAAACGCG CAGTTTATAC AAAACACAGG GGCATTATAA

ctl-1 promoter + coding region

5251 TTTTAAATC AAATTTTACA TTTTAACCAT AAAAAGCTCT TTCAAAGTG  
AAAAATTTAG TTTAAAGTGT AAAATTGGTA TTTTGTGAGA AAGTTTTCAC

ctl-1 promoter + coding region

5301 TAATTTTCTA CGCAAAAATG CCGTTCGGAT GAAAAATTAC TTTTGAAAAA  
ATTAAAAGAT GCGTTTTTAC GGCAAGCCTA CTTTTTAATG AAAACTTTTT

ctl-1 promoter + coding region

5351 CAAACTCGAA ACTACGGTAC GCAAAAAAGT ACATCGGTGT TTGCACATAA  
GTTTGAGCTT TGATGCCATG CGTTTTTICA TGTAGCCACA AACGTGTATT

ctl-1 promoter + coding region

5401 GTGAAAACAA TGTTGTTTTT TTGTAATTAA AATCGATTAA TTTTTTTTCC  
CACTTTTGTT ACAACAAAAA AACATTAATT TTAGCTAATT AAAAAAAGG

ctl-1 promoter + coding region

Fig. 17 Continued

```

=====
5451 CGGAAACAA AAACGTTTC AGCGTGGATT TCTATTGTTT CTGCGTAAA
GCCTTTTGTT TTTGCAAAAG TCGCACCTAA AGATAACAAA GAACGCATTT

                                ctl-1 promoter + coding region
=====
5501 AAAAAATTAT TTACCAATTT TAAACGATAA TTTCCACGAA TTTTCGCCAT
TTTTTTAATA AATGGTTAAA ATTTGCTATT AAAGGTGCTT AAAAGCGGTA

                                ctl-1 promoter + coding region
=====
5551 TAATCTCTCG ATTTTGTGTA TTCTTGACTC CGAGCAATCT CTCCGTTTT
ATTAGAGAGC TAAACAACCT AAGAACTGAG GCTCGTTAGA GAGGCCAAAA

                                ctl-1 promoter + coding region
=====
5601 CGCAAACGAT TATATTATTT ATTTGTTTTC CTTTTCAGTG CCGATTCTCG
GCGTTTGCTA ATATAATAAA TAAACAAAAG GAAAAGTCAC GGCTAAGAGC

ctl-1 promoter + coding region
=====
                                Exon 1
=====
5651 GAAATTCAAC AGTAAATCTT CAAAATGCCA ATGCTTCCCC ACATGGTCAA
CTTTAAGTTG TCATTTAGAA GTTTTACGGT TACGAAGGGG TGTACCAGTT

ctl-1 promoter + coding region
=====
                                Exon 1
=====
5701 TCTAAGTGAG TTTCTTTGTT ACAAATACA CGTGATGTCA GATTGTCTCA
AGATTCACCTC AAAGAAACAA TGTTTTATGT GCACTACAGT CTAACAGAGT

ctl-1 promoter + coding region
=====
5751 TTTCGGTTTG ATCTACGTAG ATCTACAAA AATGCGGGAA TTGAGCCGCA
AAAGCCAAAC TAGATGCATC TAGATGTTTT TTACGCCCTT AACTCGGCGT

ctl-1 promoter + coding region
=====
5801 GAGTTCTCAA CTGCTTTCGC ATGGTTAAGA ACGTGCGGAC GTCAAATTGT
CTCAAGAGTT GACGAAAGCG TACCAATTCT TGCACGCCTG CAGTTTAACA

ctl-1 promoter + coding region
=====
5851 TTTGGGCAAA AATTCCCGCA TTTTGTAG ATCAAACCGT AATGGGACAG
AAACCCGTTT TTAAGGGCGT AAAAAACATC TAGTTTGGA TTACCCTGTC

ctl-1 promoter + coding region
=====
                                Exon 2
=====
5901 TCTGGCACCA CGTGAÇTATA TATTTTtagc GGTCAACGAC ACAAACCCG
AGACCGTGGT GCACTGATAT ATAAAAATCG CCAGTTGCTG TGTTTTGGGC

```

Fig. 17 continued

```
ctl-1 promoter + coding region
=====
                Exon 2
=====
5951 GACCAATGGC TGAGGATCAG CTGAAAGCTT ATAGAGATAG AAATCAGGTG
    CTGGTTACCG ACTCCTAGTC GACTTTCGAA TATCTCTATC TTTAGTCCAC

ctl-1 promoter + coding region
=====
6001 AGAAAAATCA ATTTTCAGCGA TTTTCTTCGC AATTATATATA AAAACTGATT
    TCTTTTTAGT TAAAGTCGCT AAAAGAAGCG TTAAATATAT TTTTGACTAA

ctl-1 promoter + coding region
=====
                                o-GQ4
                                =====
                                Exon 3
                                =====
                                SacI
                                ~~~~~
6051 TTTCCAGGAA CCCACCTGC TCACCACATC CAATCGGAGC TCAGAAAAA
 AAAGGTCCTT GGGGTGGACG AGTGGTGTAG GTTAGCCTCG AGTCTTTTT
```

Fig. 18

# Sod-3

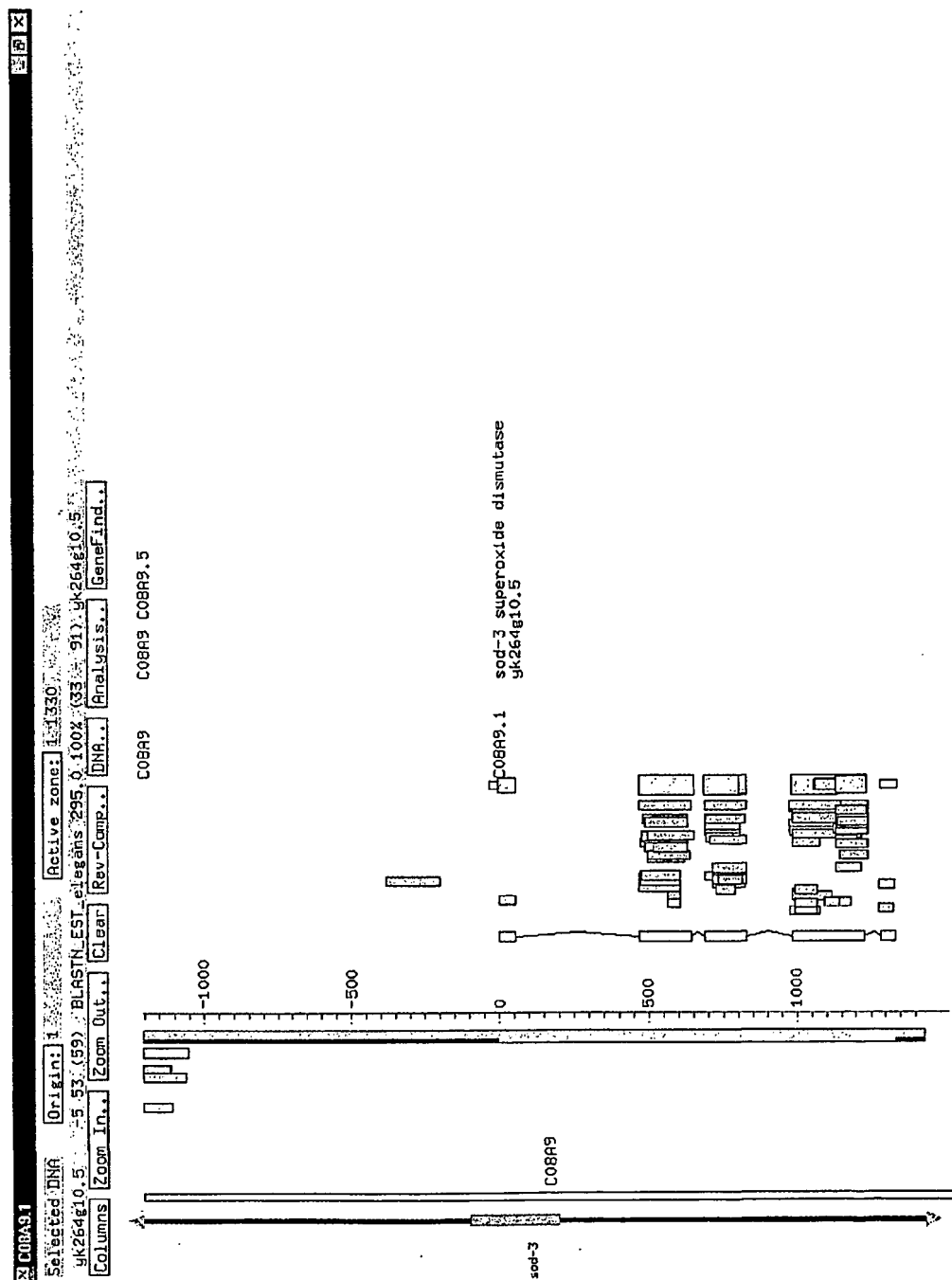


Figure 19

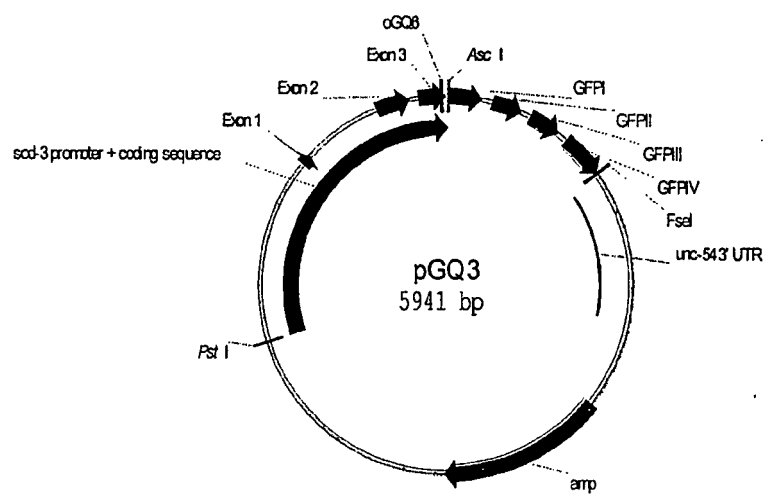


Fig. 20

## I. Predicted DNA sequence

```

oGQ6 GFPI
=====
AscI
~~~~~
1  CGCGCCATGA GTAAAGGAGA AGAACTTTTC ACTGGAGTTG TCCCAATTCT
   GCGCGGTACT CATTTCCTCT TCTTGAAAAG TGACCTCAAC AGGGTTAAGA

                                     GFPI
=====
51  TGTGAATTA GATGGTGATG TTAATGGGCA CAAATTTTCT GTCAGTGGAG
    ACAACTTAAT CTACCACTAC AATTACCCGT GTTTAAAGA CAGTCACCTC

GFPI
=====
101 AGGGTGAAGG TGATGCAACA TACGGAAAAC TTACCCTTAA ATTTATTTGC
    TCCCACCTCC ACTACGTTGT ATGCCTTTTG AATGGGAATT TAAATAAACG

GFPI
=====
151 ACTACTGGAA AACTACCTGT TCCATGGGTA AGTTTAAACA TATATATACT
    TGATGACCTT TTGATGGACA AGGTACCCAT TCAAATTTGT ATATATATGA

                                     GFPII
=====
201 AACTAACCCT GATTATTTAA ATTTTCAGCC AACACTTGTC ACTACTTTCT
    TTGATTGGGA CTAATAAATT TAAAAGTCGG TTGTGAACAG TGATGAAAGA

                                     GFPII
=====
251 GTTATGGTGT TCAATGCTTC TCGAGATACC CAGATCATAT GAAACGGCAT
    CAATACCACA AGTTACGAAG AGCTCTATGG GTCTAGTATA CTTTGCCGTA

GFPII
=====
301 GACTTTTTC AAGAGTGCCAT GCCCGAAGGT TATGTACAGG AAAGAACTAT
    CTGAAAAAGT TCTCACGGTA CGGGCTTCCA ATACATGTCC TTTCTTGATA

GFPII
=====
351 ATTTTTCAAA GATGACGGGA ACTACAAGAC ACGTAAGTTT AAACAGTTCTG
    TAAAAAGTTT CTACTGCCCT TGATGTTCTG TGCATTCAAA TTTGTCAAGC

                                     GFPIII
=====
401 GTACTAACTA ACCATACATA TTAAATTTT CAGGTGCTGA AGTCAAGTTT
    CATGATTGAT TGGTATGTAT AAATTTAAAA GTCCACGACT TCAGTTCAAA

GFPIII
=====

```

Fig. 20 continued

451 GAAGGTGATA CCCTTGTTAA TAGAATCGAG TTTAAAGGTA TTGATTTTAA  
CTTCCACTAT GGAACAATT ATCTTAGCTC AATTTTCCAT AACTAAAATT

## GFPIII

501 AGAAGATGGA AACATTCTTG GACACAAATT GGAATACAAC TATAACTCAC  
TCTTCTACCT TTGTAAGAAC CTGTGTTTAA CCTTATGTTG ATATTGAGTG

## GFPIII

551 ACAATGTATA CATCATGGCA GACAAACAAA AGAATGGAAT CAAAGTTGTA  
TGTTACATAT GTAGTACCGT CTGTTTGTTC TCTTACCTTA GTTCAACAT

## GFPIV

601 AGTTTAAACT TGGACTTACT AACTAACGGA TTATATTTAA ATTTTCAGAA  
TCAAATTTGA ACCTGAATGA TTGATTGCCT AATATAAATT TAAAAGTCTT

## GFPIV

651 CTTCAAAATT AGACACAACA TTGAAGATGG AAGCGTTCAA CTAGCAGACC  
GAAGTTTAA TCTGTGTGTG AACTTCTACC TTCGCAAGTT GATCGTCTGG

## GFPIV

701 ATTATCAACA AAATACTCCA ATTGGCGATG GCCCTGTCCT TTTACCAGAC  
TAATAGTTGT TTTATGAGGT TAACCGCTAC CGGGACAGGA AAATGGTCTG

## GFPIV

751 AACCATTACC TGTCCACACA ATCTGCCCTT TCGAAAGATC CCAACGAAAA  
TTGGTAATGG ACAGGTGTGT TAGACGGGAA AGCTTTCTAG GGTGCTTTT

## GFPIV

801 GAGAGACCAC ATGGTCCTTC TTGAGTTTGT AACAGCTGCT GGGATTACAC  
CTCTCTGGTG TACCAGGAAG AACTCAAACA TTGTCGACGA CCCTAATGTG

## GFPIV

## FseI

851 ATGGCATGGA TGAATATAC AAATAGGGCC GGCCGAGCTC CGCATCGGCC  
TACCGTACCT ACTTGATATG TTTATCCCGG CCGGCTCGAG GCGTAGCCGG

## unc-54 3' UTR

901 GCTGTCAATCA GATCGCCATC TCGCGCCCGT GCCTCTGACT TCTAAGTCCA  
CGACAGTAGT CTAGCGGTAG AGCGCGGGCA CGGAGACTGA AGATTACAGT

## unc-54 3' UTR

951 ATTACTCTTC AACATCCCTA CATGCTCTTT CTCCCTGTGC TCCCACCCCC  
TAATGAGAAG TTGTAGGGAT GTACGAGAAA GAGGGACACG AGGGTGGGGG

## unc-54 3' UTR

fig. 20 continued

```
=====
1001 TATTTTGTGTT ATTATCAAAA AAACCTCTTC TTAATTTCTT TGTTTTTTAG
    ATAAAAACAA TAATAGTTTT TTTGAAGAAG AATTAAAGAA ACAAAAAATC

                                unc-54 3' UTR
=====
1051 CTTCTTTTAA GTCACCTCTA ACAATGAAAT TGTGTAGATT CAAAAATAGA
    GAAGAAAATT CAGTGGAGAT TGTTACTTTA ACACATCTAA GTTTTATCT

                                unc-54 3' UTR
=====
1101 ATTAATTCGT AATAAAAAGT CGAAAAAAT TGTGCTCCCT CCCCCATTA
    TAATTAAGCA TTATTTTCA GCTTTTTTTA ACACGAGGA GGGGGTAAT

                                unc-54 3' UTR
=====
1151 ATAATAATTC TATCCCAAAA TCTACACAAT GTTCTGTGTA CACTTCTTAT
    TATTATTAAG ATAGGGTTTT AGATGTGTTA CAAGACACAT GTGAAGAATA

                                unc-54 3' UTR
=====
1201 GTTTTTTTTA CTTCTGATAA ATTTTTTTTG AAACATCATA GAAAAACCG
    CAAAAAAT GAAGACTATT TAAAAAAAC TTTGTAGTAT CTTTTTGGC

                                unc-54 3' UTR
=====
1251 CACACAAAAT ACCTTATCAT ATGTTACGTT TCAGTTTATG ACCGCAATTT
    GTGTGTTTTA TGGAATAGTA TACAATGCAA AGTCAAATAC TGGCGTTAA

    unc-54 3' UTR
=====
1301 TTATTTCTTC GCACGTCTGG GCCTCTCATG ACGTCAAATC ATGCTCATCG
    AATAAAGAAG CGTGCAGACC CGGAGAGTAC TGCAGTTTAG TACGAGTAGC

    unc-54 3' UTR
=====
1351 TGAAAAAGTT TTGGAGTATT TTTGGAATTT TTCAATCAAG TGAAAGTTTA
    ACTTTTTCAA AACCTCATAA AAACCTTAA AAGTTAGTC ACTTCAAAT

    unc-54 3' UTR
=====
1401 TGAAATTAAT TTTCTGCTT TTGCTTTTTG GGGGTTTCCC CTATTGTTG
    ACTTTAATTA AAAGGACGAA AACGAAAAAC CCCCAAGGG GATAACAAAC

    unc-54 3' UTR
=====
1451 TCAAGAGTTT CGAGGACGGC GTTTTCTTG CTAAATCAC AAGTATTGAT
    AGTTCTCAA GTCCTGCCG CAAAAAGAAC GATTTTAGTG TTCATACTA

    unc-54 3' UTR
=====
1501 GAGCACGATG CAAGAAAGAT CGGAAGAAG TTTGGGTTTG AGGCTCAGTG
    CTCGTGCTAC GTTCTTCTA GCCTTCTTCC AAACCAAAAC TCCGAGTCAC
```

Fig. 20 continued

unc-54 3' UTR  
=====

1551 GAAGGTGAGT AGAAGTTGAT AATTTGAAAG TGGAGTAGTG TCTATGGGGT  
CTTCCACTCA TCTTCAACTA TTAAACTTTC ACCTCATCAC AGATACCCCA

unc-54 3' UTR  
=====

1601 TTTTGCCTTA AATGACAGAA TACATTCCCA ATATACCAA CATAACTGTT  
AAAACGGAAT TTAAGTCTT ATGTAAGGGT TATATGGTTT GTATTGACAA

unc-54 3' UTR  
=

1651 TCCTACTAGT CGGCCGTACG GGCCCTTTCG TCTCGCGCGT TTCGGTGATG  
AGGATGATCA GCCGGCATGC CCGGGAAAGC AGAGCGCGCA AAGCCACTAC

1701 ACGGTGAAAA CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTTGT  
TGCCACTTTT GGAGACTGTG TACGTCGAGG GCCTCTGCCA GTGTCGAACA

1751 CTGTAAGCGG ATGCCGGGAG CAGACAAGCC CGTCAGGGCG CGTCAGCGGG  
GACATTGCGC TACGGCCCTC GTCTGTTTCG GCAGTCCCGC GCAGTCGCCC

1801 TGTTCGGCGG TGTCGGGGCT GGCTTAACTA TGCGGCATCA GAGCAGATTG  
ACAACCGCCC ACAGCCCCGA CCGAATTGAT ACGCCGTAGT CTCGTCTAAC

1851 TACTGAGAGT GCACCATATG CGGTGTGAAA TACCGCACAG ATGCGTAAGG  
ATGACTCTCA CGTGGTATAC GCCACACTTT ATGGCGTGTC TACGCATTCC

1901 AGAAAATACC GCATCAGGCG GCCTTAAGGG CCTCGTGATA CGCCTATTTT  
TCTTTTATGG CGTAGTCCGC CGGAATTCCC GGAGCACTAT GCGGATAAAA

1951 TATAGGTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT  
ATATCCAATT ACAGTACTAT TATTACCAA GAATCTGCAG TCCACCGTGA

2001 TTTCGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA  
AAAGCCCCTT TACACGCGCC TTGGGGATAA ACAAATAAAA AGATTATGT

2051 TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT  
AAGTTTATAC ATAGGCGAGT ACTCTGTTAT TGGGACTATT TACGAAGTTA

amp  
=====

2101 AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT  
TTATAACTTT TTCCTTCTCA TACTCATAAG TTGTAAAGGC ACAGCGGGAA

amp  
=====

2151 ATTCCCTTTT TTGCGGCATT TTGCCTTCCT GTTTTGTCTC ACCCAGAAAC  
TAAGGGAAAA AACGCCGTAA AACGGAAGGA CAAAACGAG TGGGTCTTTG

amp  
=====

2201 GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT  
CGACCACTTT CATTTTCTAC GACTTCTAGT CAACCCACGT GCTCACCCAA

Fig. 20 continued

amp  
=====

2251 ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC  
TGTAGCTTGA CCTAGAGTTG TCGCCATTCT AGGAACTCTC AAAAGCGGGG

amp  
=====

2301 GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC  
CTTCTTGCAA AAGGTTACTA CTCGTGAAAA TTTCAAGACG ATACACCGCG

amp  
=====

2351 GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC  
CCATAATAGG GCATAACTGC GGCCCGTTCT CGTTGAGCCA GCGGCGTATG

amp  
=====

2401 ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC AGAAAAGCAT  
TGATAAGAGT CTTACTGAAC CAACTCATGA GTGGTCAGTG TCTTTTCGTA

amp  
=====

2451 CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT  
GAATGCCTAC CGTACTGTCA TTCTCTTAAT ACGTCACGAC GGTATTGGTA

amp  
=====

2501 GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA  
CTCACTATTG TGACGCCGGT TGAATGAAGA CTGTTGCTAG CCTCCTGGCT

amp  
=====

2551 AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT  
TCCTCGATTG GCGAAAAAAC GTGTTGTACC CCCTAGTACA TTGAGCGGAA

amp  
=====

2601 GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA  
CTAGCAACCC TTGGCCTCGA CTTACTTCGG TATGGTTTGC TGCTGCACT

amp  
=====

2651 CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAAA CTATTAAGT  
GTGGTGTAC GGACATCGTT ACCGTTGTTG CAACGCGTTT GATAATTGAC

amp  
=====

2701 GCGAACTACT TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG  
CGCTTGATGA ATGAGATCGA AGGCCGTTG TTAATTATCT GACCTACCTC

amp  
=====

2751 GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG  
CGCCTATTTC AACGTCCTGG TGAAGACGCG AGCCGGGAAG GCCGACCGAC

Fig. 20 continued

```
amp
=====
2801  GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA
      CAAATAACGA CTATTTAGAC CTCGGCCACT CGCACCAGCA GCGCCATAGT

amp
=====
2851  TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC
      AACGTCGTGA CCCCAGTCTA CCATTCGGGA GGGCATAGCA TCAATAGATG

amp
=====
2901  ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA
      TGCTGCCCCCT CAGTCCGTG ATACCTACTT GCTTTATCTG TCTAGCGACT

amp
=====
2951  GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT
      CTATCCACGG AGTGACTAAT TCGTAACCAT TGACAGTCTG GTTCAAATGA

3001  CATATATACT TTAGATTGAT TTAAAACTTC ATTTTAAATT TAAAAGGATC
      GTATATATGA AATCTAATAA AATTTTGAAG TAAAAATTAA ATTTTCCTAG

3051  TAGGTGAAGA TCCTTTTGA TAATCTCATG ACCAAAATCC CTTAACGTGA
      ATCCACTTCT AGGAAAAACT ATTAGAGTAC TGGTTTTAGG GAATTGCACT

3101  GTTTTCGTTC CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT
      CAAAAGCAAG GTGACTCGCA GTCTGGGGCA TCTTTTCTAG TTTCTTAGAA

3151  CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA
      GAACTCTAGG AAAAAAGAC GCGCATTAGA CGACGAACGT TTGTTTTTTT

3201  CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT
      GGTGGCGATG GTCGCCACCA AACAAACGGC CTAGTTCTCG ATGGTTGAGA

3251  TTTTCCGAAG GTAAGTGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC
      AAAAGGCTTC CATTGACCGA AGTCGTCTCG CGTCTATGGT TTATGACAGG

3301  TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG
      AAGATCACAT CGGCATCAAT CCGGTGGTGA AGTTCTTGAG ACATCGTGCC

3351  CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG
      GGATGTATGG AGCGAGACGA TTAGGACAAT GGTCAACGAC GACGGTCACC

3401  CGATAAGTCG TGTCTTACCG GGTTGGACTC AAGACGATAG TTACCGGATA
      GCTATTCAGC ACAGAATGGC CCAACCTGAG TTCTGCTATC AATGGCCTAT

3451  AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACA GCCCAGCTTG
      TCCGCGTCGC CAGCCCGACT TGCCCCCAA GCACGTGTGT CGGGTCGAAC

3501  GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCATTGAGA
      CTCGCTTGCT GGATGTGGCT TGAATCTATG GATGTCGCAC TCGTAATCTCT
```

fig. 70 continued

3551 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG  
TTCGCGGTGC GAAGGGCTTC CCTCTTTCCG CCTGTCCATA GGCCATTTCG

3601 GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC  
CGTCCCAGCC TTGTCTCTC GCGTGCTCCC TCGAAGGTCC CCCTTTGCGG

3651 TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG  
ACCATAGAAA TATCAGGACA GCCCAAAGCG GTGGAGACTG AACTCGCAGC

3701 ATTTTGTGA TGCTCGTCAG GGGGCGGAG CCTATGAAA AACGCCAGCA  
TAAAAACACT ACGAGCAGTC CCCC GCCTC GGATACCTTT TTGCGGTCGT

3751 ACGCGGCCTT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG  
TGCGCCGGAA AAATGCCAAG GACCGGAAAA CGACCGGAAA ACGAGTGTAC

3801 TTCTTTCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT  
AAGAAAGGAC GCAATAGGGG ACTAAGACAC CTATTGGCAT AATGGCGGAA

3851 TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT  
ACTCACTCGA CTATGGCGAG CGGCGTCGGC TTGCTGGCTC GCGTCGCTCA

3901 CAGTGAGCGA GGAAGCGGAA GAGCGCCCAA TACGCAAACC GCCTCTCCCC  
GTCACCTCGT CCTTCGCCTT CTCGCGGGTT ATGCGTTTGG CGGAGAGGGG

3951 GCGCGTTGGC CGATTCAATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG  
CGCGCAACCG GCTAAGTAAT TACGTCGACC GTGCTGTCCA AAGGGGTGAC

4001 GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CTCACTCATT  
CTTTCGCCCG TCACTCGCGT TCGGTTAATT AACTCAATC GAGTGAGTAA

4051 AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA  
TCCGTGGGGT CCGAAATGTG AAATACGAAG GCCGAGCATA CAACACACCT

4101 ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA  
TAACACTCGC CTATTGTAA AGTGTGTCCT TTGTCGATAC TGGTACTAAT

sod-3 promoter + coding sequence

PstI

4151 CGCCAAGCTT GCATGCCTGC AGTGATTCAG AGAGGTTGAG AATTATTTTC  
GCGGTTTCGAA CGTACGGACG TACTAAGTC TCTCCAATC TTAATAAAAG

sod-3 promoter + coding sequence

4201 AAAAACATTC AATGTTTTCC CTTGGAGTGA CTATGCAAAT ATGAAAATGT  
TTTTTGTAAG TTACAAAAGG GAACCTCACT GATACGTTTA TACTTTTACA

sod-3 promoter + coding sequence

4251 TTTCACAAAA TATTTGGATG CCCTGATAAA AAGTAGGTGA AATTTTCGCG  
AAAGGTTTTT ATAAACCTAC GGGACTATTT TTCATCCACT TTAAAGCGTC

sod-3 promoter + coding sequence

fig. 20 continued

```

=====
4301  GGGAACATCA TATTAAAATG TTGAATTTT AGAAGAAATG GAAATGTTTG
      CCCTTG TAGT ATAATTTTAC AACTTAAAA TCTTCTTAC CTTACAAAC

                        sod-3 promoter + coding sequence
=====
4351  TCGGTGGTAT GCTCGAATAT TTGAGATATT ATATATTTAC TGTTAAATCC
      AGCCACCATA CGAGCTTATA AACTCTATAA TATATAAATG ACAATTTAGG

                        sod-3 promoter + coding sequence
=====
4401  GAAATTTTGG ACAAACGGAA AAAATTTGTG TCGAAATACT ACATTTTCGA
      CTTTAAAAAC TGTTGCCTT TTTTAAACAC AGCTTTATGA TGTAAAAGCT

                        sod-3 promoter + coding sequence
=====
4451  TAACACAAAG GTACTTCCAT AACACTTATA AAAACTGTTT GACTATCTTA
      ATTGTGTTTC CATGAAGGTA TTGTGAATAT TTTTGACAAA CTGATAGAAT

                        sod-3 promoter + coding sequence
=====
4501  TTTTCAGGAAA AAAAAATCCA AGAATAAACA TTTTTCAGAA TTTGAACTTT
      AAAGTCCTTT TTTTTCAGGT TCTTATTTGT AAAAAGTCCTT AAACCTTGAA

                        sod-3 promoter + coding sequence
=====
4551  CTAATGGCTG ATTAATAAAA CAAAGTTATA CAACTATTCA AAGCAGTTGC
      GATTACCGAC TAATTATTTT GTTCAATAT GTTGATAAGT TTCGTCAACG

                        sod-3 promoter + coding sequence
=====
4601  TCAATCTGGC ATTTTCTTGT GTTTTTTTTT GAATATTTCA TCAGCAAGAT
      AGTTAGACCG TAAAAGAACA CAAAAAATAA CTTATAAAGT AGTCGTTCTA

                        sod-3 promoter + coding sequence
=====
4651  GTTGATAATT TTGTGTTAAT TCTAATTGTT TTCTACAATT TTTCAAACCG
      CAACTATTAA AACACAATTA AGATTACAA AAGATGTTAA AAAGTTTGGC

                        sod-3 promoter + coding sequence
=====
4701  AAAATTGACC TTTGACTTTG TTTACTTTGT TCTCGTGGGT TAACTGTTCA
      TTTTAACTGG AAACGAAAC AAATGAAACA AGAGCACCCA ATGACAAGT

                        sod-3 promoter + coding sequence
=====
4751  CTGATTTCTA TTGCTGTTGA TGAGGTCTTT GATCAAATTT GTATTGTTTT
      GACTAAAGAT AACGACAAC ACTCCAGAAA CTAGTTTAAA CATAACAAAA

                        sod-3 promoter + coding sequence
=====
4801  TATACTGCAT ATTGCTTCAA TTCTAAATCA TCTAATATAT TGTCAAACAA
      ATATGACGTA TAACGAAGTT AAGATTTAGT AGATTATATA ACAGTTTGTT

```

Fig. 20 continued

```

                                sod-3 promoter + coding sequence
=====
4851  CTCTCTGTTT TTTTTTTCAT TCAAACTTC TGCAAAACG TTCTCTTAAC
      GAAGAACAAA AAAAAAGTA AGTTTGAAG ACGTTTTTGC AAGAGAATTG

                                sod-3 promoter + coding sequence
=====
4901  AAAGGTTTCAC ACAACAATC TCCTCTCCAT CTCTTTCTCT CAACAACAAT
      TTTCCAAGTG TGTTGTTGAG AGGAGAGGTA GAGAAAGAGA GTTGTTGTTA

                                sod-3 promoter + coding sequence
=====
4951  GTGCTGGCCT TGCATGTTG CCAGTGC GGG TTGTTTACGC GTTTTCAAGA
      CACGACCGGA ACGTACAAAC GGTCACGCC AACAAATGCG CAAAAGTTCT

                                sod-3 promoter + coding sequence
=====
5001  TTTTGGTCT CCTATCTAAC GTCCCGAAAT GCATTTTTTC CTTTCATTG
      AAAAACCAGA GGATAGATTG CAGGGCTTTA CGTAAAAAAG GAAAGTAAAC

                                sod-3 promoter + coding sequence
=====
5051  GTTTTTTCT GTTCGAGAAA AGTGACCGTT TGTCAAATCT TCTAATTTC
      CAAAAAAGA CAAGCTCTTT TCACTGGCAA ACAGTTTAGA AGATTAAAAG

                                sod-3 promoter + coding sequence
=====
                                Exon 1
=====
5101  AGTGAATAAA ATGCTGCAAT CTACTGCTCG CACTGCTTCA AAGCTTGTT
      TCACTTATTT TACGACGTTA GATGACGAGC GTGACGAAGT TTCGAACAAG

                                sod-3 promoter + coding sequence
=====
                                Exon 1
=====
5151  AACCGGTTGC GGGGTAAGTC AAAATGAAAT TTTCGTTTAA AAATTGGTTT
      TTGGCCAACG CCCCATTCAG TTTTACTTTA AAAGCAAATT TTTAACCAAA

                                sod-3 promoter + coding sequence
=====
5201  TTTTGGTAT TATAGATAAA ACTTATACCA AAACAAAACA TATTTAGAAA
      AAAAACCATA ATATCTATTT TGAATATGGT TTGTTTTTGT ATAAATCTTT

                                sod-3 promoter + coding sequence
=====
5251  AACTTTAATA GAGAATAATT GTTAAATAAT TAATTTTTGC AAGCTCCTTT
      TTGAAATTAT CTCTTATTAA CAAATTATTA ATTAAAAACG TTCGAGGAAA

                                sod-3 promoter + coding sequence
=====
5301  TAAATTAAGA CATCTAAAAC AGTTTTCAGC TTGATTGTTT TAATGGTTTA
      ATTTAATTCT GTAGATTTTG TCAAAAGTCG AACTAACAAA ATTACCAAT

```

Fig. 20 continued

sod-3 promoter + coding sequence

5351 GAAAGCAATA TTTGTATTTT GTGTTAACT GAAAATATCT AGGAAATACT  
CTTTCGTTAT AACATAAAA CACAATTGA CTTTATAGA TCCTTTATGA

sod-3 promoter + coding sequence

5401 ACTTTTAAAA TATTTGAAAC TTGAAATTTT AAAATTCCAA ATAATTTTAC  
TGAAAATTTT ATAACTTTG AACTTTAAAA TTTTAAGGTT TATTAAAATG

sod-3 promoter + coding sequence

5451 TCATTTCCCTA AAGTGTTTGA GTATTTGTAT CCTGTGCTGA CACCGAAATG  
AGTAAAGGAT TTCACAACT CATAAACATA GGACACGACT GTGGCTTTAC

sod-3 promoter + coding sequence

5501 TTCTCAATTT TGGAAAAAAA AGATTTTTAT CCGTATCTTC AGTCTTACAA  
AAGAGTTAAA ACCTTTTTTT TCTAAAAATA GGCATAGAAG TCAGAATGTT

sod-3 promoter + coding sequence

Exon 2

5551 TTTTTTTCAC CTTTTTTTTC ATTTCAGAGT TCTCGCCGTC CGCTCCAAGC  
AAAAAAAGTG GAAAAAAG TAAAGTCTCA AGAGCGGCAG GCGAGGTTGC

sod-3 promoter + coding sequence

Exon 2

5601 ACACTCTCCC AGATCTCCCA TTCGACTATG CAGATTGGA ACCTGTAATC  
TGTGAGAGGG TCTAGAGGGT AAGCTGATAC GTCTAAACCT TGGACATTAG

sod-3 promoter + coding sequence

Exon 2

5651 AGCCATGAAA TCATGCAGCT TCATCATCAA AAGCATCATG CCACCTACGT  
TCGGTACTTT AGTACGTCGA AGTAGTAGTT TTCGTAGTAC GGTGGATGCA

sod-3 promoter + coding sequence

Exon 2

5701 GAACAATCTC AATCAGATCG AGGAGAACT TCACGAGGCT GTTTCGAAAG  
CTTGTTAGAG TTAGTCTAGC TCCTCTTGA AGTGCTCCGA CAAAGCTTTC

sod-3 promoter + coding sequence

Exon 3

5751 GTTTTTTAAT CAGAAGATTT TGAAATGAAT TTTTTTTTGT GTATATAGGG  
CAAAAAATTA GTCTTCTAAA ACTTTACTTA AAAAAAATCATATATCCC

Fig. 20 Continued

sod-3 promoter + coding sequence

=====

Exon 3

=====

5801 AATCTAAAAG AAGCAATTGC TCTCCAACCA GCGCTGAAAT TCAATGGTGG  
TTAGATTTTC TTCGTTAACG AGAGGTTGGT CGCGACTTA AGTTACCACC

sod-3 promoter + coding sequence

=====

Exon 3

=====

5851 TGGACACATC AATCATCTTA TCTTCTGGAC CAACTTGGCT AAGGATGGTG  
ACCTGTGTAG TTAGTAAGAT AGAAGACCTG GTTGAACCGA TTCCTACCAC

oGQ6

=====

sod-3 promoter + coding sequence

=====

Exon 3

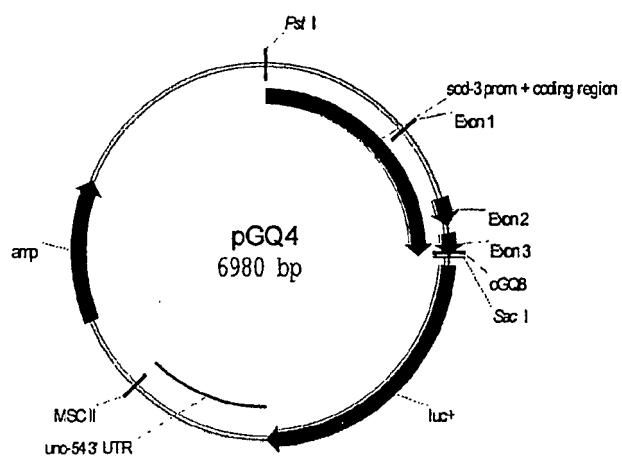
=====

AscI

~~

5901 GAGAACCTTC AAAGGAGCTG ATGGACACTA TTAAGGCTTG G  
CTCTTGGAAG TTCCTCGAC TACCTGTGAT AATCCGAAC C

Figure 21



```

=====
sod-3 prom. + coding region
=====
PstI
~
1 GTGATTCAGA GAGGTTGAGA ATTATTTTCA AAAACATTCA ATGTTTCCCG
  CACTAAGTCT CTCCAACCTCT TAATAAAAGT TTTTGTAAGT TACAAAAGGG

sod-3 prom. + coding region
=====
51 TTGGAGTGAC TATGCAAATA TGAAATGTT TTCCAAAAAT ATTTGGATGC
  AACCTCACTG ATACGTTTAT ACTTTTACAA AAGGTTTTTA TAAACCTACG

sod-3 prom. + coding region
=====
101 CCTGATAAAA AGTAGGTGAA ATTTCCGAGG GGAACATCAT ATTAAAATGT
  GGACTATTTT TCATCCACTT TAAAGCGTCC CCTGTAGTA TAATTTTACA

sod-3 prom. + coding region
=====
151 TGAATTTTTA GAAGAAATGG AAATGTTTGT CGGTGGTATG CTCGAATATT
  ACTTAAAAAT CTTCTTTACC TTTACAAACA GCCACCATAC GAGCTTATAA

sod-3 prom. + coding region
=====
201 TGAGATATTA TATATTTACT GTTAAATCCG AAATTTTGA CAAACGGAAA
  ACTCTATAAT ATATAAATGA CAATTTAGGC TTAAAAAACT GTTTGCCTTT

sod-3 prom. + coding region
=====
251 AAA'TTTGTGT CGAAATACTA CATT'TPCGAT AACACAAAGG TACTTCCATA
  TTTAAACACA GCTTTATGAT GTAAAGCTA TTGTGTTTCC ATGAAGGTAT

sod-3 prom. + coding region
=====
301 ACACTTATAA AAACTGTTTG ACTATCTTAT TTCAGGAAAA AAAAATCCAA
  TGTGAATATT TTTGACAAAC TGATAGAATA AAGTCCTTTT TTTTATAGTT

sod-3 prom. + coding region
=====
351 GAATAAACAT TTTTCAGAAT TTGAACTTTC TAATGGCTGA TTAATAAAAC
  CTTATTTGTA AAAAGTCTTA AACTTGAAAG ATTACCGACT AATTATTTTG

sod-3 prom. + coding region
=====
401 AAAGTTATAC AACTATTCAA AGCAGTTGCT CAATCTGGCA TTTTCTTGTC
  TTTCAATATG TTGATAAGTT TCGTCAACGA GTTAGACCGT AAAAGAACAC

sod-3 prom. + coding region
=====

```

Fig. 22 continued

451 TTTTTTTTGG AATATTTTCAT CAGCAAGATG TTGATAATTT TGTGTTAATT  
 AAAAAAAAAAC TTATAAAGTA GTCGTTCTAC AACTATTAAA ACACAATTAA

sod-3 prom. + coding region  
 =====

501 CTAATTGTTT TCTACAATTT TTCAAACCGA AAATTGACCT TTGACTTTGT  
 GATTAACAAA AGATGTTAAA AAGTTTGGCT TTTAACTGGA AACTGAAACA

sod-3 prom. + coding region  
 =====

551 TTACTTTGTT CTCGTGGGTT AACTGTTTAC TGATTTCTAT TGCTGTTGAT  
 AATGAAACAA GAGCACCCAA TTGACAAGTG ACTAAAGATA ACGACAACATA

sod-3 prom. + coding region  
 =====

601 GAGGTCTTTG ATCAAATTTG TATTGTTTTT AACTGTCATA TTGCTTCAAT  
 CTCAGAAAC TAGTTTAAAC ATAACAAAAA TATGACGTAT AACGAAGTTA

sod-3 prom. + coding region  
 =====

651 TCTAAATCAT CTAATATATT GTCAAACAAC TTCTTGTTTT TTTTTCATT  
 AGATTAGTA GATTATATAA CAGTTTGTG AAGAACAAAA AAAAAAGTAA

sod-3 prom. + coding region  
 =====

701 CAAACTTCT GCAAAAACGT TCTCTTAACA AAGGTTTACA CAACAACTCT  
 GTTTTGAAGA CGTTTTTGCA AGAGATTGT TTCCAAGTGT GTTGTGAGA

sod-3 prom. + coding region  
 =====

751 CCTCTCCATC TCTTTCTCTC AACAAACATG TGCTGGCCTT GCATGTTTGC  
 GGAGAGGTAG AGAAAGAGAG TTGTTGTTAC ACGACCGGAA CGTACAAACG

sod-3 prom. + coding region  
 =====

801 CAGTGCGGGT TGTTTACGCG TTTTCAAGAT TTTGGTCTC CTATCTAACG  
 GTCACGCCCA ACAATGCGC AAAAGTTCTA AAAACCAGAG GATAGATTGC

sod-3 prom. + coding region  
 =====

851 TCCCGAAATG CATTTTTTCC TTTCATTTGG TTTTCTCTG TTCGAGAAAA  
 AGGGCTTTAC GTAAAAAAGG AAAGTAAACC AAAAAAAGAC AAGCTCTTTT

sod-3 prom. + coding region  
 =====

Exon 1  
 =====

901 GTGACCGTTT GTCAAATCTT CTAATTTTCA GTGAATAAAA TGCTGCAATC  
 CACTGGCAAA CAGTTTAGAA GATTAAAAGT CACTTATTTT ACGACGTTAG

sod-3 prom. + coding region  
 =====

Exon 1  
 =====

Fig. 22 Continued

951 TACTGCTCGC ACTGCTTCAA AGCTTGTTCA ACCGGTTGCG GGGTAAGTCA  
ATGACGAGCG TGACGAAGTT TCGAACAACT TGGCCAACGC CCCATTCACT

sod-3 prom. + coding region  
=====

1001 AAATGAAATT TTCGTTTAAA AATTGGTTTT TTTGGTATT ATAGATAAAA  
TTTACTTTAA AAGCAAATTT TTAACCAAAA AAAACCATAA TATCTATTTT

sod-3 prom. + coding region  
=====

1051 CTTATACCAA AACAAAACAT ATTTAGAAAA ACTTTAATAG AGAATAATTG  
GAATATGGTT TTGTTTTGTA TAAATCTTTT TGAAATTATC TCTTATTAAC

sod-3 prom. + coding region  
=====

1101 TTTAATAATT AATTTTTGCA AGCTCCTTTT AAATTAAGAC ATCTAAAACA  
AAATTTATTAA TTAAAAACGT TCGAGGAAAA TTAATTCTG TAGATTTTGT

sod-3 prom. + coding region  
=====

1151 GTTTCAGCT TGATTGTTT AATGGTTTAG AAAGCAATAT TTGTATTTG  
CAAAAGTCGA ACTAACAAAA TTACCAATC TTTCGTTATA AACATAAAAC

sod-3 prom. + coding region  
=====

1201 TGTTAACTG AAAATATCTA GGAAATACTA CTTTTAAAT ATTTGAACT  
ACAATTTGAC TTTTATAGAT CCTTTATGAT GAAAATTTTA TAACTTTGA

sod-3 prom. + coding region  
=====

1251 TGAAATTTTA AAATTCCAAA TAATTTTACT CATTTCTTAA AGTGTGAG  
ACTTTAAAT TTTAAGGTTT ATTAAATGA GTAAAGGATT TCACAACTC

sod-3 prom. + coding region  
=====

1301 TATTTGTATC CTGTGCTGAC ACCGAAATGT TCTCAATTTT GGAAAAAAA  
ATAAACATAG GACACGACTG TGGCTTTACA AGAGTTAAAA CCTTTTTTTT

sod-3 prom. + coding region  
=====

1351 GATTTTTATC CGTATCTTCA GTCTTACAAT TTTTTTCACC TTTTTTTTCA  
CTAAAAATAG GCATAGAAGT CAGAATGTTA AAAAAAGTGG AAAAAAAGT

sod-3 prom. + coding region  
=====

Exon 2  
=====

1401 TTTCAGAGTT CTCGCCGTCC GCTCCAAGCA CACTCTCCCA GATCTCCCAT  
AAAGTCTCAA GAGCGGCAGG CGAGGTTTCGT GTGAGAGGGT CTAGAGGGTA

sod-3 prom. + coding region  
=====

Exon 2  
=====

fig. 22 continued

1451 TCGACTATGC AGATTGGAA CCTGTAATCA GCCATGAAAT CATGCAGCTT  
AGCTGATACG TCTAACCTT GGACATTAGT CCGTACTTTA GTACGTCGAA

sod-3 prom. + coding region  
=====

Exon 2  
=====

1501 CATCATCAAA AGCATCATGC CACCTACGTG AACAACTCA ATCAGATCGA  
GTAGTAGTTT TCGTAGTACG GTGGATGCAC TTGTTAGAGT TAGTCTAGCT

sod-3 prom. + coding region  
=====

Exon 2  
=====

1551 GGAGAACTT CACGAGGCTG TTTCGAAAGG TTTTAAATC AGAAGATTTT  
CCTCTTTGAA GTGCTCCGAC AAAGCTTTCC AAAAATTAG TCTCTAAAA

sod-3 prom. + coding region  
=====

Exon 3  
=====

1601 GAAATGAATT TTTTTTTGG TATATAGGGA ATCTAAAAGA AGCAATTGCT  
CTTACTTAA AAAAAAACC ATATATCCCT TAGATTTTCT TCGTTAACGA

sod-3 prom. + coding region  
=====

Exon 3  
=====

1651 CTCCAACCAG CGCTGAAATT CAATGGTGGT GGACACATCA ATCATCTAT  
GAGGTTGGTC GCGACTTAA GTTACCACCA CCTGTGTAGT TAGTAAGATA

oGQ8  
=====

sod-3 prom. + coding region  
=====

Exon 3  
=====

1701 CTTCTGGACC AACTTGGCTA AGGATGGTGG AGAACCTTCA AAGGAGCTGA  
GAAGACCTGG TTGAACCGAT TCCTACCACC TCTTGAAGT TTCCTCGACT

oGQ8  
=====

sod-3 prom. + coding region  
=====

Exon 3  
=====

SacI  
~~~~~

1751 TGGACACTAT TAAGCCGAGC TCAGAAAAA TGACTGCTCC AAAGAAGAAG
ACCTGTGATA ATTCGGCTCG AGTCTTTTTT ACTGACGAGG TTTCTTCTTC

luc+
=====

1801 CGTAAGGTAC CGGTAGAAAA AATGGAAGAC GCCAAAAACA TAAAGAAAGG

Fig. 22 continued

	GCATTCCATG	GCCATCTTTT	TTACCTTCTG	CGGTTTTTGT	ATTCTCTTCC	luc+
	=====					
1851	CCCGGCGCCA	TTCTATCCGC	TGGAAGATGG	AACCGCTGGA	GAGCAACTGC	
	GGGCCGCGGT	AAGATAGGCG	ACCTTCTACC	TTGGCGACCT	CTCGTTGACG	
	=====					
	luc+					
	=====					
1901	ATAAGGCTAT	GAAGAGATAC	GCCCTGGTTC	CTGGAACAAT	TGCTTTTACA	
	TATTCCGATA	CTTCTCTATG	CGGGACCAAG	GACCTTGTTA	ACGAAAATGT	
	=====					
	luc+					
	=====					
1951	GATGCACATA	TCGAGGTGGA	CATCACTTAC	GCTGAGTACT	TCGAAATGTC	
	CTACGTGTAT	AGCTCCACCT	GTAGTGAATG	CGACTCATGA	AGCTTTACAG	
	=====					
	luc+					
	=====					
2001	CGTTCCGTTG	GCAGAAGCTA	TGAAACGATA	TGGGCTGAAT	ACAAATCACA	
	GCAAGCCAAC	CGTCTTCGAT	ACTTTGCTAT	ACCCGACTTA	TGTTTAGTGT	
	=====					
	luc+					
	=====					
2051	GAATCGTCGT	ATGCAGTGAA	AACTCTCTTC	AATTCTTTAT	GCCGGTGTTG	
	CTTAGCAGCA	TACGTCACTT	TTGAGAGAAG	TTAAGAAATA	CGGCCACAAC	
	=====					
	luc+					
	=====					
2101	GGCGCGTTAT	TTATCGGAGT	TGCAGTTGCG	CCCGCGAACG	ACATTTATAA	
	CCGCGCAATA	AATAGCCTCA	ACGTCAACGC	GGGCGCTTGC	TGTAATATT	
	=====					
	luc+					
	=====					
2151	TGAACGTGAA	TTGCTCAACA	GTATGGGCAT	TTCGCAGCCT	ACCGTGGTGT	
	ACTTGCACTT	AACGAGTTGT	CATACCCGTA	AAGCGTCGGA	TGGCACCACA	
	=====					
	luc+					
	=====					
2201	TCGTTTCCAA	AAAGGGGTTG	CAAAAAATTT	TGAACGTGCA	AAAAAAGCTC	
	AGCAAAGGTT	TTTCCCAAC	GTTTTTTAAA	ACTTGACAGT	TTTTTTCGAG	
	=====					
	luc+					
	=====					
2251	CCAATCATCC	AAAAAATTAT	TATCATGGAT	TCTAAAACGG	ATTACCAGGG	
	GGTTAGTAGG	TTTTTTAATA	ATAGTACCTA	AGATTTTGCC	TAATGGTCCC	
	=====					
	luc+					
	=====					
2301	ATTTCAAGTCG	ATGTACACGT	TCGTACATC	TCATCTACCT	CCCGGTTTTTA	
	TAAAGTCAGC	TACATGTGCA	AGCAGTGTAG	AGTAGATGGA	GGGCCAAAAT	
	=====					
	luc+					
	=====					

fig. 22 Continued

2351 ATGAATACGA TTTTGTGCCA GAGTCCTTCG ATAGGGACAA GACAATTGCA
TACTTATGCT AAAACACGGT CTCAGGAAGC TATCCCTGTT CTGTTAACGT
luc+
=====

2401 CTGATCATGA ACTCCTCTGG ATCTACTGGT CTGCCTAAAG GTGTCGCTCT
GACTAGTACT TGAGGAGACC TAGATGACCA GACGGATTTC CACAGCGAGA
luc+
=====

2451 GCCTCATAGA ACTGCCTGCG TGAGATTCTC GCATGCCAGA GATCCTATTT
CGGAGTATCT TGACGGACGC ACTCTAAGAG CGTACGGTCT CTAGGATAAA
luc+
=====

2501 TTGGCAATCA AATCATTCCG GATACTGCGA TTTTAAGTGT TGTTCATTTC
AACCGTTAGT TTAGTAAGGC CTATGACGCT AAAATTCACA ACAAGGTAAG
luc+
=====

2551 CATCACGGTT TTGGAATGTT TACTACACTC GGATATTTGA TATGTGGATT
GTAGTGCCAA AACCTTACAA ATGATGTGAG CCTATAAACT ATACACCTAA
luc+
=====

2601 TCGAGTCGTC TTAATGTATA GATTGAAGA AGAGCTGTTT CTGAGGAGCC
AGCTCAGCAG AATTACATAT CTAACTTCT TCTCGACAAA GACTCCTCGG
luc+
=====

2651 TTCAGGATTA CAAGATTCAA AGTGCGCTGC TGGTGCCAAC CCTATTCTCC
AAGTCCTAAT GTTCTAAGTT TCACGCGACG ACCACGGTTG GGATAAGAGG
luc+
=====

2701 TTCTTCGCCA AAAGCACTCT GATTGACAAA TACGATTAT CTAATTACA
AAGAAGCGGT TTTCGTGAGA CTAAGTGTGTT ATGCTAAATA GATTAAATGT
luc+
=====

2751 CGAAATTGCT TCTGGTGGCG CTCCCCTCTC TAAGGAAGTC GGGGAAGCGG
GCTTTAACGA AGACCACCGC GAGGGGAGAG ATTCCTTCAG CCCCTTCGCC
luc+
=====

2801 TTGCCAAGAG GTTCCATCTG CCAGGTATCA GGCAAGGATA TGGGCTCACT
AACGGTTCTC CAAGGTAGAC GGTCCATAGT CCGTTCCTAT ACCCGAGTGA
luc+
=====

2851 GAGACTACAT CAGCTATTCT GATTACACCC GAGGGGGATG ATAAACCGGG
CTCTGATGTA GTCGATAAGA CTAATGTGGG CTCCCCCTAC TATTGGCCCC
luc+

Fig. 22 continued

```
=====
2901  CGCGGTCGGT AAAGTTGTTC CATTTTTTGA AGCGAAGGTT GTGGATCTGG
      GCGCCAGCCA TTCAACAAG GTAAAAAACT TCGCTTCCAA CACCTAGACC

      luc+
      =====
2951  ATACCGGGAA AACGCTGGGC GTTAATCAAA GAGGCGAACT GTGTGTGAGA
      TATGCCCCTT TTGCGACCCG CAATTAGTTT CTCCGCTTGA CACACACTCT

      luc+
      =====
3001  GGTCCATAGA TTATGTCCGG TTATGTAAAC AATCCGGAAG CGACCAACGC
      CCAGGATACT AATACAGGCC AATACATTG TTAGGCCTTC GCTGGTTGCG

      luc+
      =====
3051  CTTGATTGAC AAGGATGGAT GGCTACATTC TGGAGACATA GCTTACTGGG
      GAACTAACTG TTCCTACCTA CCGATGTAAG ACCTCTGTAT CGAATGACCC

      luc+
      =====
3101  ACGAAGACGA ACACTTCTTC ATCGTTGACC GCCTGAAGTC TCTGATTAAg
      TGCTTCTGCT TGTGAAGAAG TAGCAACTGG CGGACTTCAG AGACTAATTC

      luc+
      =====
3151  TACAAAGGCT ATCAGGTGGC TCCCGCTGAA TTGGAATCCA TCTTGCTCCA
      ATGTTTCCGA TAGTCCACCG AGGGCGACTT AACCTTAGGT AGAACGAGGT

      luc+
      =====
3201  ACACCCCAAC ATCTTCGACG CAGGTGTCGC AGGTCTTCCC GACGATGACG
      TGTGGGGTTG TAGAAGCTGC GTCCACAGCG TCCAGAAGGG CTGCTACTGC

      luc+
      =====
3251  CCGGTGAACT TCCCGCCGCC GTTGTTGTTT TGGAGCACGG AAAGACGATG
      GGCCACTTGA AGGGCGGCGG CAACAACAAA ACCTCGTGCC TTTCTGCTAC

      luc+
      =====
3301  ACGGAAAAAG AGATCGTGGA TTACGTCGCC AGTCAAGTAA CAACCGCGAA
      TGCCTTTTTC TCTAGCACCT AATGCAGCGG TCAGTTCATT GTTGGCGCTT

      luc+
      =====
3351  AAAGTTGCGC GGAGGAGTTG TGTTTGTTGA CGAAGTACCG AAAGGTCTTA
      TTTCAACGCG CCTCCTCAAC ACAAACACCT GCTTCATGGC TTTCCAGAAT

      luc+
      =====
3401  CCGGAAAACT CGACGCAAGA AAAATCAGAG AGATCCTCAT AAAGGCCAAG
      GGCCTTTTGA GCTGCGTTCT TTTTAGTCTC TCTAGGAGTA TTTCCGGTTC
```

Fig. 22 continued

```

luc+                                unc-54 3' UTR
=====
3451 AAGGGCGGAA AGATCGCCGT GTAATTCTAG GAATTCCAAC TGAGCGCCGG
    TTCCCGCCTT TCTAGCGGCA CATTAAGATC CTTAAGGTTG ACTCGCGGCC

                                unc-54 3' UTR
=====
3501 TCGCTACCAT TACCAACTTG TCTGGTGTCA AAAATAATAG GGGCCGCTGT
    AGCGATGGTA ATGGTTGAAC AGACCACAGT TTTTATTATC CCCGGCGACA

                                unc-54 3' UTR
=====
3551 CATCAGAGTA AGTTTAAACT GAGTTCTACT AACTAACGAG TAATATTTAA
    GTAGTCTCAT TCAAATTTGA CTCAAGATGA TTGATTGCTC ATTATAAATT

                                unc-54 3' UTR
=====
3601 ATTTTCAGCA TCTCGCGCCC GTGCCTCTGA CTTCTAAGTC CAATTACTCT
    TAAAAGTCGT AGAGCGCGGG CACGGAGACT GAAGATTCAG GTTAATGAGA

                                unc-54 3' UTR
=====
3651 TCAACATCCC TACATGCTCT TTCTCCCTGT GCTCCCACCC CCTATTTTGT
    AGTTGTAGGG ATGTACGAGA AAGAGGGACA CGAGGGTGGG GGATAAAAAC

                                unc-54 3' UTR
=====
3701 TTATTATCAA AAAAATTCT TCTTAATTC TTTGTTTTTT AGCTTCTTTT
    AATAATAGTT TTTTGAAGA AGAATTAAAG AAACAAAAAA TCGAAGAAAA

                                unc-54 3' UTR
=====
3751 AAGTCACCTC TAACATGAA ATTGTGTAGA TTCAAAAATA GAATTAATTC
    TTCAGTGGAG ATTGTTACTT TAACACATCT AAGTTTTTAT CTTAATTAAG

                                unc-54 3' UTR
=====
3801 GTAATAAAAA GTCGAAAAA ATTGTGCTCC CTCCCCCAT TAATAATAAT
    CATTATTTTT CAGCTTTTTT TAACACGAGG GAGGGGGGTA ATTATTATTA

                                unc-54 3' UTR
=====
3851 TCTATCCCAA AATCTACACA ATGTTCTGTG TACACTTCTT ATGTTTTTTT
    AGATAGGGTT TTAGATGTGT TACAAGACAC ATGTGAAGAA TACAAAAAAA

                                unc-54 3' UTR
=====
3901 TACTTCTGAT AAATTTTTTT TGAAACATCA TAGAAAAAAC CGCACACAAA
    ATGAAGACTA TTTAAAAAAA ACTTTGTAGT ATCTTTTTTG GCGTGTGTTT

                                unc-54 3' UTR
=====
3951 ATACCTTATC ATATGTTACG TTTCAGTTTA TGACCGCAAT TTTTATTCTT
    TATGGAATAG TATACAATGC AAAGTCAAAT ACTGGCGTTA AAAATAAAGA

```

Fig. 22 continued

```
unc-54 3' UTR
=====
4001 TCGCACGTCT GGGCCTCTCA TGACGTCAAA TCATGCTCAT CGTGAAAAAG
    AGCGTGCAGA CCCGGAGAGT ACTGCAGTTT AGTACGAGTA GCACTTTTTC

unc-54 3' UTR
=====
4051 TTTTGGAGTA TTTTGGGAAT TTTTCAATCA AGTGAAAGTT TATGAAATTA
    AAAACCTCAT AAAACCTTA AAAAGTTAGT TCACTTTCAA ATACTTTAAT

unc-54 3' UTR
=====
4101 ATTTTCCTGC TTTTGCTTTT TGGGGGTTTC CCCTATTGTT TGTCAAGAGT
    TAAAAGGACG AAAACGAAAA ACCCCCAAAG GGCATAACAA ACAGTTCTCA

unc-54 3' UTR
=====
4151 TTCGAGGACG GCGTTTTTCT TGCTAAAATC ACAAGTATTG ATGAGCACGA
    AAGCTCCTGC CGCAAAAAGA ACGATTTTAG TGTCATAAC TACTCGTGCT

unc-54 3' UTR
=====
4201 TGCAAGAAAG ATCGGAAGAA GGTTTGGGTT TGAGGCTCAG TGGAAGGTGA
    ACGTTCTTTC TAGCCTTCTT CCAAACCCAA ACTCCGAGTC ACCTTCCACT

unc-54 3' UTR
=====
4251 GTAGAAGTTG ATAATTGAA AGTGGAGTAG TGCTATGGG GTTTTGCCT
    CATCTTCAAC TATTAACTT TCACCTCATC ACAGATACCC CAAAAACGGA

unc-54 3' UTR
=====
4301 TAAATGACAG AATACATTCC CAATATACCA AACATAACTG TTTCTACTA
    ATTTACTGTC TTATGTAAGG GTTATATGGT TTGTATTGAC AAAGGATGAT

MSC II
=====
4351 GTCGGCCGTA CGGGCCCTTT CGTCTCGCGC GTTTCGGTGA TGACGGTGAA
    CAGCCGGCAT GCCCGGGAAA GCAGAGCGCG CAAAGCCACT ACTGCCACTT

4401 AACCTCTGAC ACATGCAGCT CCCGGAGACG GTCACAGCTT GTCTGTAAGC
    TTGGAGACTG TGTACGTCGA GGGCCTCTGC CAGTGTCGAA CAGACATTGC

4451 GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG GGTGTGGCG
    CCTACGGGCC TCGTCTGTTC GGGCAGTCCC GCGCAGTCGC CCACAACCGC

4501 GGTGTCCGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT TGTACTGAGA
    CCACAGCCCC GACCGAATTG ATACGCCGTA GTCTCGTCTA ACATGACTCT

4551 GTGCACCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA
    CACGTGGTAT ACGCCACACT TTATGGCGTG TCTACGCATT CCTCTTTTAT

4601 CCGCATCAGG CGGCCTTAAG GGCCTCGTGA TACGCCTATT TTTATAGGTT
```

Fig. 22 continued

GGCGTAGTCC GCCGGAATTC CCGGAGCACT ATGCGGATAA AAATATCCAA

4651 AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGTGGCA CTTTTCGGGG
TTACAGTACT ATTATTACCA AAGAATCTGC AGTCCACCGT GAAAAGCCCC

4701 AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA
TTTACACGCG CCTTGGGGAT AAACAAATAA AAAGATTTAT GTAAGTTTAT

4751 TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA
ACATAGGCGA GTACTCTGTT ATTGGGACTA TTTACGAAGT TATTATAACT

amp
=====

4801 AAAAGGAAGA GTATGAGTAT TCAACATTTT CGTGTCGCCC TTATTCCCTT
TTTTCTTCT CATACTCATA AGTTGTAAAG GCACAGCGGG AATAAGGGAA

amp
=====

4851 TTTTTCGGCA TTTTGCCTTC CTGTTTTCG TCACCCAGAA ACGCTGGTGA
AAAACGCCGT AAAACGGAAG GACAAAAACG AGTGGGTCTT TCGGACCACT

amp
=====

4901 AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG TTACATCGAA
TTCATTTTCT ACGACTTCTA GTCAACCCAC GTGCTCAGCC AATGTAGCTT

amp
=====

4951 CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG
GACCTAGAGT TGTCGCCATT CTAGGAACTC TCAAAGCGG GGCTTCTTGC

amp
=====

5001 TTTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC GCGGTATTAT
AAAAGGTTAC TACTCGTGAA AATTTCAGAA CGATACACCG CGCCATAATA

amp
=====

5051 CCCGTATTGA CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT
GGGCATAACT GCGGCCCGTT CTCGTTGAGC CAGCGGCGTA TGTGATAAGA

amp
=====

5101 CAGAATGACT TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA
GTCTTACTGA ACCAACTCAT GAGTGGTCAG TGTCTTTTCG TAGAATGCCT

amp
=====

5151 TGGCATGACA GTAAGAGAAT TATGCAGTGC TGCCATAACC ATGAGTGATA
ACCGTACTGT CATTCTCTTA ATACGTCACG ACGGTATTGG TACTCACTAT

amp
=====

5201 ACACTGCGGC CAACTTACTT CTGACAACGA TCGGAGGACC GAAGGAGCTA

Fig. 22 continued

TGTGACGCCG GTTGAATGAA GACTGTTGCT AGCCTCCTGG CTCCTCGAT

amp

=====

5251 ACCGCTTTT TGCACAACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG
TGGCGAAAAA ACSTGTTGTA CCCCCTAGTA CATTGAGCGG AACTAGCAAC

amp

=====

5301 GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCACGA
CCTTGGCCTC GACTTACTTC GGTATGGTTT GCTGCTCGCA CTGTGGTGC

amp

=====

5351 TGCCTGTAGC AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA
ACGGACATCG TTACCGTTGT TGCAACGCGT TTGATAATTG ACCGCTTGAT

amp

=====

5401 CTTACTCTAG CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA
GAATGAGATC GAAGGGCCGT TGTTAATTAT CTGACCTACC TCCGCCTATT

amp

=====

5451 AGTTGCAGGA CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG
TCAACGTCCT GGTGAAGACG CGAGCCGGGA AGGCCGACCG ACCAAATAAC

amp

=====

5501 CTGATAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT CATTGCAGCA
GACTATTTAG ACCTCGGCCA CTCGCACCCA GAGCGCCATA GTAACGTCGT

amp

=====

5551 CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG
GACCCCGGTC TACCATTTCG GAGGGCATAG CATCAATAGA TGTGCTGCCC

amp

=====

5601 GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG
CTCAGTCCGT TGATACCTAC TTGCTTTATC TGTCTAGCGA CTCTATCCAC

amp

=====

5651 CCTCACTGAT TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA
GGAGTGAATA ATTCGTAACC ATTGACAGTC TGGTTCAAAT GAGTATATAT

5701 CTTTAGATTG ATTTAAACT TCATTTTAA TTTAAAAGGA TCTAGGTGAA
GAAATCTAAC TAAATTTTGA AGTAAAAATT AAATTTTCCT AGATCCACTT

5751 GATCCTTTTT GATAATCTCA TGACCAAAAT CCCTTAACGT GAGTTTTCGT
CTAGGAAAAA CTATTAGAGT ACTGGTTTTA GGAATTGCA CTCAAAAGCA

5801 TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT

fig 22. *continued*

```

AGGTGACTCG CAGTCTGGGG CATCTTTTCT AGTTTCCTAG AAGAACTCTA

5851 CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT
GGAAAAAAG ACGCGCATTA GACGACGAAC GTTTGTTTTT TTGGTGCGCA

5901 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTCCTGA
TGGTCGCCAC CAAACAAACG GCCTAGTTCT CGATGGTTGA GAAAAAGGCT

5951 AGGTAAGTGG CTTCAGCAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG
TCCATTGACC GAAGTCGTCT CGCGTCTATG GTTTATGACA GGAAGATCAC

6001 TAGCCGTAGT TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCTACATA
ATCGGCATCA ATCCGGTGGT GAAGTTCTTG AGACATCGTG GCGGATGTAT

6051 CCTCGCTCTG CTAATCCTGT TACCAGTGGC TGCTGCCAGT GGCATAAGT
GGAGCGAGAC GATTAGGACA ATGGTCACCG ACGACGGTCA CCCTATTCA

6101 CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG
GCACAGAATG GCCCAACCTG AGTTCTGCTA TCAATGGCCT ATTCCGCGTC

6151 CGGTGCGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC
GCCAGCCCGA CTTGCCCCCC AAGCACGTGT GTCGGGTCGA ACCTCGCTTG

6201 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCATTGA GAAAGCGCCA
CTGGATGTGG CTTGACTCTA TGGATGTCGC ACTCGTAACT CTTTCGCGGT

6251 CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC
CGGAAGGGCT TCCCTCTTTC CGCCTGTCCA TAGGCCATTG GCCGTCCCAG

6301 GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGGAACG CCTGGTATCT
CCTTGTCCTC TCGCGTGCTC CCTCGAAGGT CCCCCTTTCG GGACCATAGA

6351 TTATAGTCCT GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTTGT
AATATCAGGA CAGCCCAAAG CGGTGAGAC TGAACGCA GCTAAAAACA

6401 GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG CAACGCGGCC
CTACGAGCAG TCCCCCGGCC TCGGATACCT TTTTGCGGTC GTTGCGCGCG

6451 TTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTCTTTCC
AAAAATGCCA AGGACCGGAA AACGACCGGA AAACGAGTGT ACAAGAAAGG

6501 TGCATTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG
ACGCAATAGG GACTAAGAC ACCTATTGGC ATAATGGCGG AACTCACTC

6551 CTGATACCGC TCGCCGACG CGAACGACCG AGCGCAGCGA GTCAGTGAGC
GACTATGGCG AGCGGCGTCG GCTTGCTGGC TCGCGTCGCT CAGTCACTCG

6601 GAGGAAGCGG AAGAGCGCCC AATACGCAA CCGCCTCTCC CCGCGCGTTG
CTCCTTCGCC TTCTCGCGGG TTATGCGTTT GGCGGAGAGG GGCGCGCAAC

6651 GCCGATTCAT TAATGCAGCT GGCACGACAG GTTCCCGAC TGGAAAGCGG
CGGCTAAGTA ATTACGTCGA CCGTGCTGTC CAAAGGGCTG ACCTTTCGCC

6701 GCAGTGAGCG CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC

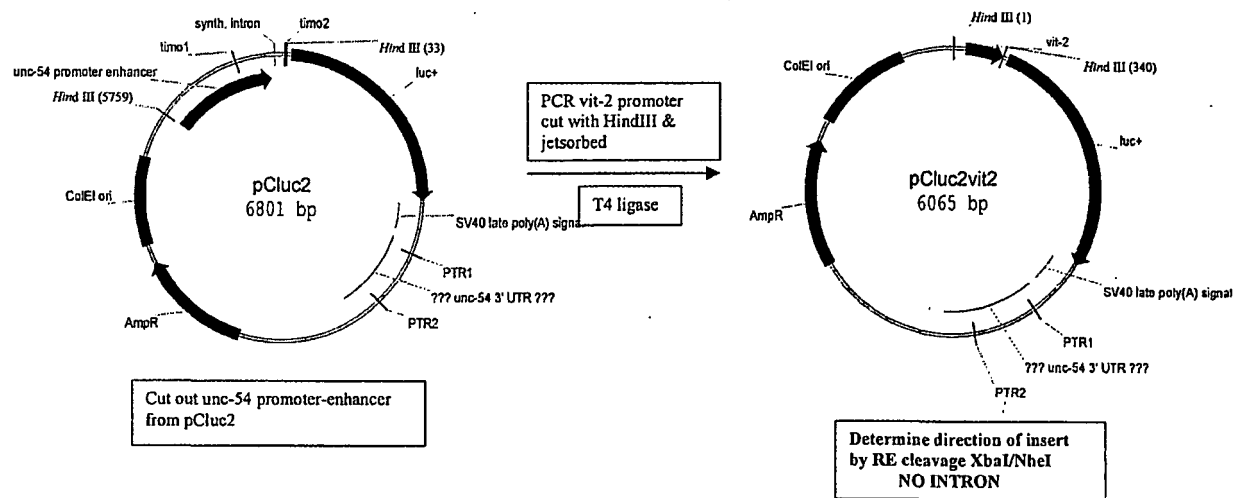
```

Fig-22 continued

```
CGTCACTCGC GTTGCGTTAA TTAACTCAA TCGAGTGAGT AATCCGTGGG
6751 CAGGCTTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG
    GTCCGAAATG TGAAATACGA AGGCCGAGCA TACAACACAC CTTAACTCTC
6801 CGGATAACAA TTTCACACAG GAAACAGCTA TGACCATGAT TACGCCAAGC
    GCCTATTGTT AAAGTGTGTC CTTGTGCGAT ACTGGTACTA ATGCGGTTTC
6851 TGTAAGTTTA AACATGATCT TACTAACTAA CTATTCTCAT TTAAATTTTC
    ACATTCAAAT TTGTACTAGA ATGATTGATT GATAAGAGTA AATTAAAAAG
6901 AGAGCTTAAA AATGGCTGAA ATCACTCACA ACGATGGATA CGCTAACAAAC
    TCTCGAATTT TTACCGACTT TAGTGAGTGT TGCTACCTAT GCGATTGTTG

                                PstI
                                ~~~~~
6951 TTGGAATGA AATAAGCTTG CATGCCTGCA
    AACCTTTACT TTATTCGAAC GTACGGACGT
```

Figure 23



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
13 December 2001 (13.12.2001)

PCT

(10) International Publication Number
WO 01/93669 A3

(51) International Patent Classification⁷: **G01N 33/50**

(21) International Application Number: **PCT/IB01/01199**

(22) International Filing Date: **8 June 2001 (08.06.2001)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:
0014009.5 **8 June 2000 (08.06.2000) GB**

(71) Applicant (for all designated States except US): **DEV-GEN NV [BE/BE];** Technologiepark 9, B-9052 Zwijnaarde (BE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **FEICHTINGER, Richard [AT/BE];** Walpoortstraat 9, B-9000 Gent (BE). **BOGAERT, Thierry [BE/BE];** Wolvendreef 26g, 8500 Kortrijk (BE).

(74) Agents: **BALDOCK, Sharon, Claire et al.;** Boulton, Watt, Tennant, Verulam Gardens, 70 Gray's Inn Road, London WC1X 8BT (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)
- of inventorship (Rule 4.17(iv)) for US only

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
10 May 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/93669 A3

(54) Title: **COMPOUND SCREENS RELATING TO INSULIN DEFICIENCY OR INSULIN RESISTANCE**

(57) Abstract: The invention is concerned with use of the model organism *C. elegans* as a research tool to screen for compounds active in insulin signalling. In particular, the invention relates to improved screening methods based on release of *C. elegans* from the dauer larval state.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 01/01199

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, WPI Data, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 51351 A (GEN HOSPITAL CORP) 19 November 1998 (1998-11-19) cited in the application claims 1-8	1-62
A	<p>--- GEMS DAVID ET AL: "Two pleiotropic classes of daf-2 mutation affect larval arrest, adult behavior, reproduction and longevity in Caenorhabditis elegans." GENETICS, vol. 150, no. 1, 1998, pages 129-155, XP002191748 ISSN: 0016-6731 cited in the application the whole document --- -/--</p>	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

Z document member of the same patent family

Date of the actual completion of the international search

28 February 2002

Date of mailing of the international search report

15/03/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040. Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Niemann, F

INTERNATIONAL SEARCH REPORT

II International Application No

PCT/IB 01/01199

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GIL E B ET AL: "REGULATION OF THE INSULIN-LIKE DEVELOPMENTAL PATHWAY OF CAENORHABDITIS ELEGANS BY A HOMOLOG OF THE PTEN TUMOR SUPPRESSOR GENE" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 96, March 1999 (1999-03), pages 2925-2930, XP002926980 ISSN: 0027-8424 abstract ----	
A	KIMURA K D ET AL: "DAF-2, AN INSULIN RECEPTOR-LIKE GENE THAT REGULATES LONGEVITY AND DIAPAUSE IN CAENORHABDITIS ELEGANS" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 277, 15 August 1997 (1997-08-15), pages 942-946, XP002910188 ISSN: 0036-8075 cited in the application the whole document ----	
P,X	WO 00 33068 A (GEN HOSPITAL CORP) 8 June 2000 (2000-06-08) claims 1-14 -----	1,16

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 01/01199

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9851351	A	19-11-1998	US 6225120 B1	01-05-2001
			AU 7494198 A	08-12-1998
			EP 1019092 A1	19-07-2000
			PL 336858 A1	17-07-2000
			WO 9851351 A1	19-11-1998
			HU 0002199 A2	28-09-2000
<hr/>				
WO 0033068	A	08-06-2000	US 2001029617 A1	11-10-2001
			AU 1749600 A	19-06-2000
			EP 1163515 A1	19-12-2001
			WO 0033068 A1	08-06-2000
<hr/>				

Amendment Transmittal Letter
Application No. 11/191,863

* IF HIGHEST NUMBER PREVIOUSLY PAID FOR IS 20 OR LESS, WRITE "20" IN COLUMN 3
** IF HIGHEST NUMBER PREVIOUSLY PAID FOR IS 3 OR LESS, WRITE "3" IN COLUMN 3
*** PAY THIS FEE ONLY WHEN MULTIPLE DEPENDENT CLAIMS ARE ADDED FOR THE FIRST
TIME

_____ Attached is our check for \$ to pay the fees calculated above.
_____ A Petition for Extension of Time and the required fee are enclosed.
_____ Other enclosures:

The Commissioner is hereby authorized to charge any fees under 37 CFR 1.16 and 1.17 which may be required by or to give effect to this paper to Deposit Account No. 03-1728. Please show our docket number with any charge or credit to our Deposit Account. **A copy of this letter is enclosed.**

Respectfully submitted,

CHRISTIE PARKER & HALE, LLP

By _____

Constantine Marantidis
Reg. No. 39,759
626/795-9900

CM/scc

SCC PAS744450.1-*07/5/07 11:54 AM